

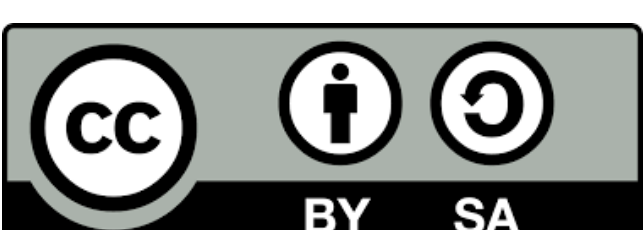


Zestaw startowy dla etycznej
sztucznej inteligencji

Przewodnik po fundamentach etycznej
sztucznej inteligencji

www.valuesai.eu

2025



Co-funded by
the European Union

Finansowane przez Unię Europejską. Wyrażone poglądy i opinie są jednak wyłącznie poglądami i opiniami autora/autorów i niekoniecznie odzwierciedlają stanowisko Unii Europejskiej ani Europejskiej Agencji Wykonawczej ds. Edukacji i Kultury (EACEA). Ani Unia Europejska, ani EACEA nie ponoszą za nie odpowiedzialności.

Podręcznik podstaw etyki sztucznej inteligencji © 2025 autorstwa VALUES jest udostępniany na podstawie licencji CC BY 4.0.

Przedstawiamy projekt VALUES

Projekt VALUES to innowacyjna inicjatywa, której celem jest wzmocnienie pozycji młodych ludzi w całej Europie poprzez wyposażenie ich w niezbędne umiejętności oraz zrozumienie etyki, kluczowe do poruszania się w dynamicznie rozwijającym się świecie sztucznej inteligencji (AI). W miarę jak technologie AI stają się integralną częścią codziennego życia, niezwykle istotne jest, aby młodzież nie tylko zdobywała wiedzę techniczną, ale także rozwijała silną świadomość etycznych aspektów związanych z AI. Głównym celem projektu jest wyposażenie młodych ludzi, zwłaszcza tych zmagających się z problemami na rynku pracy, w wiedzę i narzędzia niezbędne do krytycznej oceny technologii AI, co umożliwi im odpowiedzialne korzystanie z tych systemów. Poprzez promowanie głębszego zrozumienia etyki AI, projekt VALUES ma na celu przygotowanie młodych ludzi do roli odpowiedzialnych obywateli cyfrowych oraz etycznych innowatorów, gotowych do wniesienia pozytywnego wkładu w gospodarkę cyfrową. Oczekiwane rezultaty obejmują wzrost wiedzy na temat AI, bardziej etyczne podejście do korzystania z technologii oraz zwiększenie możliwości zatrudnienia w rozwijającym się sektorze AI.

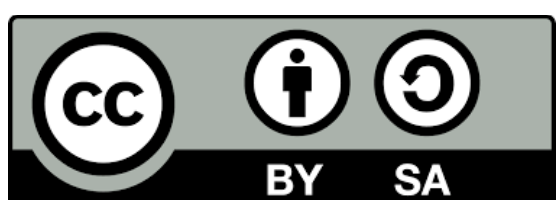
Czym jest Zestaw Startowy VALUES dla Etycznej Sztucznej Inteligencji?

Zestaw startowy VALUES dla etycznej AI stanowi Twoją bramę do odkrywania fascynującego i istotnego świata etycznej sztucznej inteligencji. Zaprojektowany z myślą o pracownikach młodzieżowych i edukatorach, ten dynamiczny zasób oferuje narzędzia do zgłębiania złożoności AI, od jej podstawowych zasad po dylematy etyczne, które ze sobą niesie. Zestaw, bogaty w przykłady z życia, praktyczne strategie oraz angażujące ćwiczenia, pomoże Ci rozpocząć

wartościowe dyskusje z młodymi ludźmi na temat wpływu AI na ich życie. Niezależnie od tego, czy dopiero zaczynasz przygodę z AI, czy wspierasz młodych ludzi w zrozumieniu jej etycznych implikacji, ten zestaw startowy zwiększy Twoją pewność siebie w nauczaniu i wzbudzi ciekawość dotyczącą odpowiedzialnego wykorzystania AI we współczesnym świecie.

Spis treści

01	WSTĘP DO SZTUCZNEJ INTELIGENCJI I ETYKI
02	PODSTAWOWE ZASADY ETYKI SZTUCZNEJ INTELIGENCJI
03	DYLEMATY ETYCZNE W DZIEDZINIE SZTUCZNEJ INTELIGENCJI
04	PRZEPROWADZANIE MŁODZIEŻY PRZEZ DYSKUSJE ETYCZNE NA TEMAT SZTUCZNEJ INTELIGENCJI
05	ZAŁĄCZNIKI



Co-funded by
the European Union

Finansowane przez Unię Europejską. Wyrażone poglądy i opinie są jednak wyłącznie poglądami i opiniami autora/autorów i niekoniecznie odzwierciedlają stanowisko Unii Europejskiej ani Europejskiej Agencji Wykonawczej ds. Edukacji i Kultury (EACEA). Ani Unia Europejska, ani EACEA nie ponoszą za nie odpowiedzialności.

Podręcznik podstaw etyki sztucznej inteligencji © 2025 autorstwa VALUES jest udostępniany na podstawie licencji CC BY 4.0.

01

WSTĘP DO SZTUCZNEJ INTELIGENCJI I ETYKI



Wprowadzenie do sztucznej inteligencji oraz etyki

Czym jest sztuczna inteligencja?

Sztuczna inteligencja (AI) to dziedzina informatyki, której celem jest opracowywanie systemów zdolnych do realizacji zadań tradycyjnie wymagających ludzkiej inteligencji. Systemy te mają na celu symulację ludzkich funkcji poznawczych, takich jak uczenie się, rozumowanie, rozwiązywanie problemów, percepcja oraz rozumienie języka. Celem AI jest nie tylko naśladowanie ludzkich działań, ale także umożliwienie maszynom samodoskonalenia się w miarę zdobywania doświadczenia i danych. W dzisiejszym świecie AI dynamicznie przekształca różne branże, a zrozumienie jej kluczowych koncepcji jest niezbędne dla pracowników młodzieżowych, edukatorów oraz młodzieży, aby mogli skutecznie poruszać się w tym zmieniającym się otoczeniu.

Przyjrzyjmy się dokładniej fundamentalnym koncepcjom sztucznej inteligencji:

1. Uczenie maszynowe (ML)

Uczenie maszynowe stanowi jedną z kluczowych dziedzin sztucznej inteligencji (AI), w której algorytmy pozwalają komputerom na samodzielne uczenie się oraz formułowanie prognoz na podstawie danych, bez potrzeby ich bezpośredniego programowania. W istocie, w uczeniu maszynowym system zwiększa swoją efektywność, analizując obszerne zbiory danych i identyfikując w nich wzorce.

Na przykład uczenie maszynowe stanowi fundament systemów rekomendacji wykorzystywanych przez platformy takie jak Netflix i YouTube. Systemy te analizują wcześniejsze działania użytkowników (takie jak historia oglądania), aby proponować filmy lub nagrania wideo, które mogą im się spodobać.

Przykład z rzeczywistego świata:

- **Rekomendacje Netflix:** Netflix stosuje uczenie maszynowe do proponowania filmów i seriali na podstawie historii oglądania, preferencji oraz ocen. Im więcej danych system zbiera o Twoim zachowaniu, tym trafniej przewidzi, co zechcesz obejrzeć następnym razem.

2. Przetwarzanie języka naturalnego (NLP)

Przetwarzanie języka naturalnego (NLP) pozwala komputerom na zrozumienie, interpretację i reakcję na język ludzki w sposób zarówno znaczący, jak i użyteczny. NLP wspiera maszyny w przetwarzaniu języka w różnych formach, takich jak język mówiony (rozpoznawanie mowy) oraz tekst pisany (analiza tekstu). NLP znajduje zastosowanie w aplikacjach takich jak chatboty, wirtualni asystenci oraz narzędzia tłumaczeniowe.

Jednym z fundamentalnych wyzwań w przetwarzaniu języka naturalnego jest nauczenie maszyn rozumienia kontekstu, tonu oraz znaczenia języka ludzkiego. Umożliwia to komputerom nie tylko interpretację słów, ale także uchwycenie ukrytego sensu zdania lub frazy.

Przykład z rzeczywistego świata:

- **Asystenci głosowi (Siri, Alexa):** Urządzenia te wykorzystują przetwarzanie języka naturalnego (NLP) do interpretacji poleceń głosowych oraz adekwatnego reagowania. Na przykład, gdy poprosisz Alexę o odtworzenie utworu, NLP umożliwi jej zrozumienie Twojej prośby i precyzyjne wykonanie zadania.



OBEJRZYJ: „Przetwarzanie języka naturalnego w pięć minut”

Wprowadzenie do sztucznej inteligencji oraz etyki

Czym jest sztuczna inteligencja?

3. Głębokie uczenie maszynowe

Uczenie głębokie to poddziedzina uczenia maszynowego, która wykorzystuje sztuczne sieci neuronowe do symulacji sposobu przetwarzania informacji przez ludzki mózg. Sieci te składają się z licznych warstw węzłów (stąd termin „głębokie”), co pozwala systemowi na naukę i podejmowanie decyzji z wysoką precyzją.

Głębokie uczenie wykazuje szczególną skuteczność w zadaniach takich jak rozpoznawanie obrazów, gdzie system uczy się identyfikować obiekty, twarze lub sceny na fotografiach. Im więcej danych przetwarza system, tym lepiej rozpoznaje wzorce i formułuje prognozy.

Przykład z rzeczywistego świata:

- **Technologia rozpoznawania twarzy:** Algorytmy głębokiego uczenia wspierają systemy rozpoznawania twarzy wykorzystywane w kamerach bezpieczeństwa, smartfonach oraz platformach mediów społecznościowych. Systemy te potrafią identyfikować osoby na zdjęciach lub filmach, co przyczynia się do zwiększenia bezpieczeństwa oraz personalizacji.

4. Wizja komputerowa

Widzenie komputerowe to obszar sztucznej inteligencji, który pozwala maszynom na interpretację i zrozumienie informacji wizualnych z otaczającego je świata. Obejmuje to zadania takie jak identyfikacja obiektów na obrazach, rozpoznawanie twarzy oraz interpretacja filmów. Widzenie komputerowe znajduje zastosowanie w wielu sektorach, w tym w opiece zdrowotnej, handlu detalicznym oraz transporcie.

Na przykład, samochody autonomiczne w znacznym stopniu wykorzystują wizję komputerową do identyfikacji przeszkód, sygnalizacji świetlnej oraz pieszych na drodze. Przetwarzając w czasie rzeczywistym obrazy z kamer zamontowanych w

pojeździe, samochód jest w stanie podejmować decyzje dotyczące swojego ruchu.

Przykład z rzeczywistego świata:

- **Samochody autonomiczne:** Firmy takie jak Tesla i Waymo stosują wizję komputerową, aby umożliwić autonomiczną jazdę. Pojazdy te wykorzystują algorytmy sztucznej inteligencji do analizy danych z kamer i czujników, co pozwala im na poruszanie się po ulicach oraz podejmowanie decyzji, takich jak zatrzymanie się na sygnalizacji świetlnej czy omijanie przeszkód.

5. Sztuczna inteligencja w codziennym życiu

Choć powyższe koncepcje mogą wydawać się techniczne, sztuczna inteligencja stała się nieodłącznym elementem naszego codziennego życia. Od mediów społecznościowych po zakupy online, technologie sztucznej inteligencji wspierają podejmowanie decyzji, poprawiają doświadczenia oraz automatyzują zadania. Oto kilka kluczowych przykładów zastosowań sztucznej inteligencji:

- **Algorytmy mediów społecznościowych:** Sztuczna inteligencja napędza algorytmy, które decydują o tym, jakie posty, reklamy i treści są wyświetlane na platformach takich jak Facebook, Instagram i Twitter. Systemy te analizują Twoje zachowanie oraz preferencje, aby generować spersonalizowane treści.
- **Zakupy online:** Detaliczni sprzedawcy, tacy jak Amazon, stosują sztuczną inteligencję do rekomendowania produktów na podstawie Twoich wcześniejszych zakupów lub historii wyszukiwania. Ten system rekomendacji wykorzystuje uczenie maszynowe, aby przewidywać, co prawdopodobnie nabędziesz w przyszłości.
- **Opieka zdrowotna:** Sztuczna inteligencja wkracza do medycyny, wspierając lekarzy w diagnozowaniu chorób, takich jak rak, prognozowaniu wyników leczenia pacjentów oraz tworzeniu spersonalizowanych planów terapeutycznych na podstawie analizy danych.

Wprowadzenie do sztucznej inteligencji oraz etyki

Rola sztucznej inteligencji w społeczeństwie.

Sztuczna inteligencja staje się coraz bardziej integralną częścią naszego codziennego życia, kształtując sposób, w jaki wchodzimy w interakcje ze światem, pracujemy oraz komunikujemy się w społeczeństwie. Nie ogranicza się jedynie do laboratoriów technologicznych czy dużych korporacji; znalazła swoje miejsce na platformach społecznościowych, w miejscu pracy, a nawet w życiu obywatelskim, wpływając i transformując różne sektory oraz obszary życia. Ta sekcja omawia rolę sztucznej inteligencji w społeczeństwie, koncentrując się na jej wpływie na życie społeczne, karierę zawodową oraz zaangażowanie młodzieży.

1. Typowe zastosowania w życiu społecznym.

Sztuczna inteligencja jest obecnie wszechobecna w życiu społecznym, a jej wpływ jest wyraźnie odczuwalny w wielu dziedzinach, zwłaszcza w mediach społecznościowych oraz technologiach rozrywkowych. Oto kilka najczęstszych sposobów, w jakie sztuczna inteligencja kształtuje doświadczenia społeczne:

Algorytmy platform społecznościowych

Platformy mediów społecznościowych, takie jak Facebook, Instagram, TikTok i Twitter, stosują algorytmy oparte na sztucznej inteligencji do generowania i personalizowania treści dla użytkowników. Algorytmy te analizują Twoje interakcje – polubienia, udostępnienia, komentarze oraz czas spędzony na postach – aby określić, które treści najprawdopodobniej przyciągną Twoją uwagę w przyszłości. Im częściej angażujesz się w określone rodzaje treści, tym bardziej sztuczna inteligencja platformy dostosowuje Twój kanał do Twoich preferencji.

Algorytmy te mają na celu zwiększenie zaangażowania użytkowników, co skutkuje dłuższym czasem spędzonym przed ekranem, a w konsekwencji większymi możliwościami kierowania do nich spersonalizowanych reklam. Proces ten rodzi jednak również pytania dotyczące prywatności, gromadzenia danych oraz potencjalnej manipulacji użytkownikami przez ukierunkowane treści, w tym dezinformację i sensację.

Przykład z rzeczywistego świata:

- **Strona „Eksploruj” na Instagramie:** Sztuczna inteligencja odpowiedzialna za stronę „Eksploruj” na Instagramie prognozuje, które posty mogą Cię zainteresować, opierając się na Twoich wcześniejszych polubieniach oraz wzorcach aktywności. Jeśli regularnie lubisz zdjęcia związane z modą, algorytm Instagrama zaprezentuje Ci więcej treści w tej dziedzinie, dostosowując Twoje doświadczenie, aby utrzymać Twoje zainteresowanie.

Deepfake'i

Deepfake to syntetyczne materiały generowane przez sztuczną inteligencję, które nakładają obrazy lub filmy na istniejące nagrania, często tworząc hiperrealistyczne, fałszywe filmy. Tego rodzaju filmy są powszechnie wykorzystywane do podszywania się pod kogoś, sprawiając wrażenie, że dana osoba powiedziała lub zrobiła coś, czego w rzeczywistości nigdy nie uczyniła. Technologia deepfake wykorzystuje techniki sztucznej inteligencji, w szczególności głębokie uczenie, do tworzenia przekonujących manipulacji, które mogą obejmować zarówno nieszkodliwą rozrywkę, jak i szkodliwe treści mające na celu oszukanie lub manipulację.

Rozwój technologii deepfake budzi istotne obawy etyczne, szczególnie gdy są one wykorzystywane do szerzenia dezinformacji, znieśławień lub manipulacji politycznych. Podważają one wiarygodność źródeł medialnych i mogą osłabić zaufanie do treści wizualnych.

Przykład z rzeczywistego świata:

- **Filmy deepfake w polityce:** Deepfake'i były wykorzystywane w kampaniach politycznych do tworzenia sfabrykowanych przemówień lub wydarzeń z udziałem osób publicznych, co może potencjalnie wpływać na wybory lub opinię publiczną. Możliwość manipulacji treściami wideo rodzi istotne pytania dotyczące tego, jak weryfikujemy informacje i jaką mamy ufność w źródła medialne.

Wprowadzenie do sztucznej inteligencji oraz etyki

Rola sztucznej inteligencji w społeczeństwie.

2. Wpływ sztucznej inteligencji na rozwój kariery

Wpływ sztucznej inteligencji na rynek pracy jest znaczący i przekształca wiele branż, generując nowe możliwości zatrudnienia, a jednocześnie powodując, że niektóre zawody stają się nieaktualne. W miarę postępu technologii związanej ze sztuczną inteligencją, zapotrzebowanie na umiejętności w tym zakresie gwałtownie rośnie, co czyni ten obszar kluczowym dla młodych ludzi wkraczających na rynek pracy.

Automatyzacja i eliminacja miejsc pracy

Jednym z najczęściej analizowanych efektów sztucznej inteligencji (AI) jest automatyzacja – zastosowanie maszyn i algorytmów do realizacji zadań tradycyjnie wykonywanych przez ludzi. Automatyzacja, mimo że zwiększa efektywność, stwarza również zagrożenie dla bezpieczeństwa zatrudnienia w różnych sektorach, szczególnie w produkcji, transporcie i wprowadzaniu danych. Roboty oraz algorytmy oparte na AI już zastępują powtarzalne zadania, które niegdyś wymagały pracy ludzkiej, co prowadzi do utraty miejsc pracy przez pracowników w tych branżach.

Przykład z rzeczywistego świata:

- **Kasy samoobsługowe:** Wiele sklepów detalicznych obecnie wdraża kasy samoobsługowe, co prowadzi do zmniejszenia zapotrzebowania na kasjerów. W miarę udoskonalania tych systemów, mogą one zastąpić liczne stanowiska w sektorze handlu detalicznego, dlatego istotne jest, aby pracownicy podnosili swoje kwalifikacje i uczyli się obsługi systemów automatycznych.

Rosnące zapotrzebowanie na kompetencje związane ze sztuczną inteligencją

Z drugiej strony, sztuczna inteligencja (AI) otwiera nowe możliwości zatrudnienia w takich dziedzinach jak analiza danych, uczenie maszynowe, robotyka oraz etyka AI. Wraz z integracją technologii AI w różnych sektorach przemysłu, rośnie zapotrzebowanie na specjalistów z wiedzą w tych obszarach. Dla młodych ludzi nabycie umiejętności w

zakresie AI i pokrewnych dziedzin może znacząco poprawić ich perspektywy zawodowe, zapewniając solidne przygotowanie do odniesienia sukcesu na przyszłym rynku pracy.

Sztuczna inteligencja nie ogranicza się jedynie do sektora technologicznego. Technologie AI znajdują zastosowanie również w takich dziedzinach jak opieka zdrowotna, prawo, finanse oraz marketing. Na przykład w opiece zdrowotnej sztuczna inteligencja jest wykorzystywana do prognozowania wyników leczenia pacjentów, wspierania diagnostyki obrazowej oraz personalizacji planów terapeutycznych, co stwarza nowe możliwości dla specjalistów w zakresie sztucznej inteligencji w tych obszarach.

Przykład z rzeczywistego świata:

- **Sztuczna inteligencja w opiece zdrowotnej:** Przedsiębiorstwa takie jak IBM Watson Health stosują sztuczną inteligencję do analizy danych medycznych oraz dostarczania analiz, które wspierają lekarzy w formułowaniu dokładniejszych diagnoz. Ten postęp wymaga wykwalifikowanych specjalistów do zarządzania i implementacji systemów sztucznej inteligencji w placówkach opieki zdrowotnej.

PRZECZYTAJ:
Wpływ sztucznej
inteligencji na
zawody, rynek
pracy i
zatrudnienie



Wprowadzenie do sztucznej inteligencji oraz etyki

Rola sztucznej inteligencji w społeczeństwie.

3. Rola sztucznej inteligencji w życiu młodzieży (aspekty społeczne, zawodowe, zaangażowanie obywatelskie)

Rola sztucznej inteligencji wykracza daleko poza sferę zawodową i media społecznościowe – ma ona również istotny wpływ na sposób, w jaki młodzi ludzie angażują się w życie swoich społeczności oraz w sprawy globalne. Rosnąca integracja sztucznej inteligencji z różnymi aspektami życia oznacza, że młodzież musi być dobrze poinformowana i wyposażona w narzędzia, aby mogła korzystać z tych technologii w sposób wartościowy, zarówno społecznie, jak i obywatelsko.

Życie społeczne oraz aktywność obywatelska

Sztuczna inteligencja wpływa na sposób, w jaki młodzi ludzie wchodzi w interakcje z otaczającym ich światem. Od algorytmów mediów społecznościowych po narzędzia oparte na sztucznej inteligencji wykorzystywane w aktywizmie, sztuczna inteligencja odgrywa kluczową rolę w kształtowaniu sposobu rozpowszechniania i konsumpcji informacji. Zrozumienie sztucznej inteligencji umożliwia młodym ludziom krytyczne podejście do treści, które napotykają w internecie, podejmowanie świadomych decyzji dotyczących swojego cyfrowego śladu oraz unikanie manipulacji, zarówno ze strony reklam, jak i dezinformacji.

Sztuczna inteligencja jest coraz częściej wykorzystywana w aktywizmie obywatelskim, gdzie wzmacnia głos młodzieży w istotnych kwestiach społecznych, takich jak zmiany klimatyczne, prawa człowieka oraz edukacja. Platformy oparte na sztucznej inteligencji mogą na przykład wspierać aktywistów w śledzeniu i udostępnianiu w czasie rzeczywistym informacji o protestach, mobilizowaniu zwolenników oraz analizowaniu trendów społecznych, które mogą być przydatne w opracowywaniu strategii aktywistycznych.

Przykład z rzeczywistego świata:

- **Sztuczna inteligencja w ruchach społecznych:**

Platformy takie jak Change.org wykorzystują sztuczną inteligencję do organizowania petycji oraz mobilizowania ludzi wokół kwestii społecznych. Sztuczna inteligencja potrafi analizować trendy w mediach społecznościowych, aby identyfikować kluczowe problemy i zwiększać świadomość na temat istotnych zagadnień.

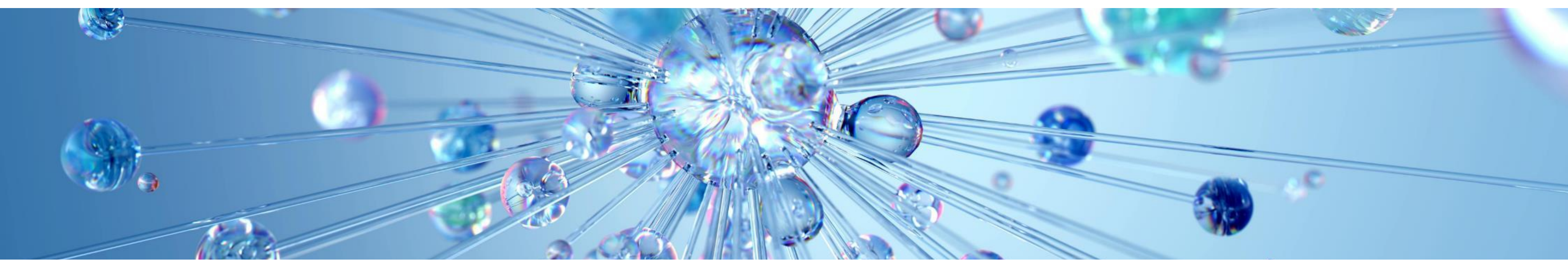
Kariera oraz edukacja

Dla młodych ludzi wchodzących na rynek pracy sztuczna inteligencja staje się kluczowym elementem wielu dziedzin. Jak wcześniej wspomniano, branże takie jak opieka zdrowotna, finanse czy rozrywka coraz bardziej polegają na technologiach sztucznej inteligencji. Zrozumienie sztucznej inteligencji pozwala młodzieży lepiej przygotować się do kariery w tych sektorach. Co więcej, umiejętności związane z AI zyskują na znaczeniu w edukacji. Szkoły i uniwersytety włączają AI do swoich programów nauczania, aby uczyć studentów nie tylko samej AI, ale także krytycznego myślenia o jej wpływie na społeczeństwo.

Pracownicy młodzieżowi oraz edukatorzy pełnią kluczową rolę w przygotowaniu przyszłych pokoleń do odpowiedzialnego i innowacyjnego korzystania ze sztucznej inteligencji. Wyposażenie młodych ludzi w wiedzę na temat sztucznej inteligencji nie ogranicza się jedynie do przekazywania umiejętności technicznych; chodzi o umożliwienie im aktywnego, świadomego i etycznego uczestnictwa w społeczeństwie, w którym sztuczna inteligencja odgrywa coraz bardziej znaczącą rolę.

Przykład z rzeczywistego świata:

- **Sztuczna inteligencja a spersonalizowane nauczanie:** W obszarze edukacji sztuczna inteligencja jest stosowana do kreowania spersonalizowanych doświadczeń edukacyjnych. Platformy takie jak Khan Academy wykorzystują sztuczną inteligencję do dostosowywania materiałów edukacyjnych w zależności od indywidualnych postępów uczniów, oferując każdemu z nich treści dostosowane do ich potrzeb.



Wniosek: Wprowadzenie do sztucznej inteligencji oraz etyki

Jak zauważyliśmy w tym rozdziale, sztuczna inteligencja to nie tylko innowacja technologiczna – to potężna siła kształtująca nasze życie społeczne, kariery oraz zaangażowanie obywatelskie. Sztuczna inteligencja jest już integralną częścią naszych codziennych doświadczeń, od algorytmów mediów społecznościowych, które formują nasze kanały, po jej wpływ na rynek pracy, generując nowe możliwości, a jednocześnie stawiając nowe wyzwania. Dla młodych ludzi sztuczna inteligencja to nie tylko koncepcja przyszłości – to rzeczywistość, która dotyka wielu aspektów ich życia i będzie odgrywać jeszcze większą rolę, gdy wejdą na rynek pracy i zaangażują się w otaczający ich świat.

Etyczne implikacje sztucznej inteligencji są głębokie i dalekosiężne. Niezależnie od tego, czy dotyczy to zapewnienia przejrzystości algorytmów, zwalczania stronniczości, czy ochrony danych osobowych, odpowiedzialność za etyczne korzystanie ze sztucznej inteligencji spoczywa na nas wszystkich. Ten rozdział wprowadza w fundamentalne koncepcje sztucznej inteligencji oraz jej podstawowe zasady etyczne, przygotowując grunt pod dalsze, bardziej zaawansowane dyskusje. Zrozumienie tych koncepcji stanowi pierwszy krok w kierunku wspierania młodych ludzi w odpowiedzialnym poruszaniu się po złożonym krajobrazie sztucznej inteligencji oraz podejmowaniu świadomych decyzji zgodnych z wytycznymi etycznymi i wartościami europejskimi.

Jako pracownicy młodzieżowi, edukatorzy i trenerzy, niezwykle istotne jest, abyśmy wyposażyli młodych ludzi zarówno w umiejętności techniczne, jak i zdolność krytycznego myślenia, niezbędne do efektywnego korzystania z technologii AI. Kolejne rozdziały zgłębią szczegółowe dylematy etyczne związane z AI oraz zaproponują praktyczne narzędzia, które wspomogą młodzież w dyskusjach na te tematy, przygotowując ich do bycia nie tylko kompetentnymi użytkownikami AI, ale także odpowiedzialnymi i etycznymi uczestnikami w świecie zdominowanym przez AI.





„Sztuczna inteligencja nie jest neutralna. To nie tylko narzędzie; to odzwierciedlenie wartości ludzi, którzy ją tworzą”.

Timnit Gebru, Badacz sztucznej inteligencji i orędownik etycznej sztucznej inteligencji



02

PODSTAWOWE ZASADY ETYKI SZTUCZNEJ INTELIGENCJI



Podstawowe zasady etyki sztucznej inteligencji

Czym jest etyka i dlaczego ma znaczenie?

Etyka to dziedzina filozofii, która bada kwestie dotyczące moralności, definiując, co jest słuszne, a co niesłuszne; dobre, a co złe; sprawiedliwe, a co niesprawiedliwe. W kontekście sztucznej inteligencji (AI) etyka odgrywa kluczową rolę, ponieważ systemy AI mogą wpływać na istotne aspekty życia ludzkiego, od interakcji społecznych po możliwości zatrudnienia, opiekę zdrowotną, a nawet wolność osobistą. W miarę jak systemy AI coraz bardziej kształtują naszą rzeczywistość, niezwykle istotne jest, aby były projektowane, rozwijane i wykorzystywane zgodnie z fundamentalnymi zasadami moralnymi, takimi jak uczciwość, sprawiedliwość i poszanowanie praw jednostki.

Etyka w sztucznej inteligencji (AI) umożliwia zrozumienie złożoności interakcji tych technologii z ludźmi oraz społeczeństwem. Technologie AI często podejmują decyzje lub formułują sugestie na podstawie ogromnych zbiorów danych, co może prowadzić do niezamierzonych konsekwencji, jeśli systemy nie zostaną starannie zaprojektowane z uwzględnieniem zasad etycznych. Na przykład algorytmy AI mogą nieumyślnie utrzymywać stereotypy, dyskryminować określone grupy lub naruszać prywatność jednostek. Aby zapobiec tym negatywnym skutkom, zasady etyczne muszą być integrowane na każdym etapie rozwoju AI, od projektowania i gromadzenia danych, po wdrażanie i użytkowanie.

Znaczenie etyki w sztucznej inteligencji (AI) jest szczególnie istotne, ponieważ systemy AI stają się coraz bardziej autonomiczne i wpływowe. W przypadku braku nadzoru etycznego, systemy AI mogą utrzymywać nierówności, pogłębiać istniejące uprzedzenia lub podejmować decyzje naruszające podstawowe prawa człowieka. Na przykład systemy oparte na AI wykorzystywane w procesach rekrutacyjnych mogą nieumyślnie faworyzować określone grupy kosztem innych, jeśli nie będą odpowiednio monitorowane pod kątem uprzedzeń. Podobnie technologie AI stosowane w monitoringu mogą naruszać prywatność, jeśli nie będą odpowiednio regulowane.

Etyka w AI jest kluczowa dla budowania zaufania między ludźmi a tymi technologiami. Gdy użytkownicy nie mają pewności co do odpowiedzialnego wykorzystania systemów AI, mogą stać się niechętni do ich wdrażania, mimo że mogą one przynieść znaczące korzyści. Z drugiej strony, jeśli systemy AI są postrzegane jako etyczne, uczciwe i transparentne, jednostki oraz społeczeństwo chętniej z nich korzystają. Wprowadzając zasady etyczne w AI, możemy tworzyć systemy, które są nie tylko skuteczne i wydajne, ale także sprawiedliwe i równe.

Ostatecznie, rozważania etyczne dotyczące sztucznej inteligencji wykraczają poza unikanie szkód; koncentrują się na zapewnieniu, że systemy sztucznej inteligencji są projektowane i wykorzystywane w sposób, który przynosi korzyści społeczeństwu. Etyka sztucznej inteligencji polega na podejmowaniu przemyślanych decyzji dotyczących zastosowania technologii dla dobra wspólnego, przy jednoczesnym poszanowaniu praw i wolności jednostki. Etyczna sztuczna inteligencja skłania nas do krytycznego myślenia o wpływie, jaki chcemy, aby te systemy miały na nasz świat, oraz o tym, jak tworzyć technologię zgodną z naszymi wartościami.

W miarę jak zmierzamy ku przyszłości, w której sztuczna inteligencja będzie odgrywać coraz większą rolę, zrozumienie i rozwiązanie kwestii etycznych staje się kluczowe dla zapewnienia, że technologie te będą służyć ludzkości w sposób uczciwy, sprawiedliwy i zgodny z naszym wspólnym moralnym kompasem.



Podstawowe zasady etyki sztucznej inteligencji

Błąd w algorytmach sztucznej inteligencji.

Systemy sztucznej inteligencji podejmują decyzje na podstawie danych, jednak gdy dane te są stronnicze, wyniki również mogą być zniekształcone. Stronniczość w sztucznej inteligencji odnosi się do niesprawiedliwych lub nierównych rezultatów, które mogą dyskryminować określone grupy. Może to wystąpić na różnych etapach, zwłaszcza podczas gromadzenia, szkolenia i zastosowania danych.

Jak kształtują się uprzedzenia w sztucznej inteligencji

Uprzedzenia występują, gdy systemy sztucznej inteligencji są szkolone na danych, które są niekompletne, niereprezentatywne lub obciążone historycznymi uprzedzeniami. Na przykład systemy rozpoznawania twarzy często mają trudności z identyfikacją kobiet oraz osób o innym kolorze skóry niż biały, ponieważ dane treningowe koncentrują się głównie na osobach białych. Podobnie, sztuczna inteligencja stosowana w procesie rekrutacji może utrzymywać uprzedzenia w decyzjach rekrutacyjnych podejmowanych w przeszłości, niesprawiedliwie dyskryminując kobiety lub grupy mniejszościowe.

Rodzaje uprzedzeń w sztucznej inteligencji.

- Błąd selekcji: występuje, gdy użyte dane nie odzwierciedlają populacji.
- Błąd etykietowania: występuje, gdy etykiety w zbiorze danych treningowych są nieprecyzyjne lub stronnicze.
- Błąd pomiaru: występuje w wyniku niedokładnego lub stronnicego gromadzenia danych.
- Błąd agregacji: występuje, gdy sztuczna inteligencja generalizuje cechy różnych grup jako jednorodne.

Rozwiązywanie problemów dotyczących uprzedzeń w sztucznej inteligencji

1. Różnorodne i reprezentatywne dane: zapewnienie, że dane szkoleniowe odzwierciedlają wszystkie grupy demograficzne.
2. Audyty stronniczości: systematyczne testowanie systemów AI w celu identyfikacji i eliminacji

stronniczości.

3. Przejrzyste algorytmy: Opracowywanie wyjaśnialnych algorytmów w celu zrozumienia i korygowania stronnicej decyzji.

4. Sprawiedliwość algorytmiczna: integracja kryteriów sprawiedliwości w fazie projektowania.

5. Nadzór ludzki: zaangażowanie ludzi w proces podejmowania decyzji, szczególnie w kontekście aplikacji o kluczowym znaczeniu.

Implikacje etyczne

Uprzedzenia w sztucznej inteligencji mogą utrzymywać nierówności społeczne, naruszać prawa człowieka oraz podważać zaufanie do technologii. Na przykład, stronnicze algorytmy rekrutacyjne mogą wykluczać wykwalifikowanych kandydatów na podstawie płci lub rasy, a stronnicze algorytmy wymiaru sprawiedliwości mogą prowadzić do niesprawiedliwych wyroków. Przejrzystość i odpowiedzialność są kluczowe dla zapobiegania tym problemom etycznym oraz budowania zaufania do systemów sztucznej inteligencji.

Wniosek

Aby tworzyć etyczne systemy sztucznej inteligencji, kluczowe jest identyfikowanie, rozumienie i łagodzenie uprzedzeń. Zapewniając uczciwość i sprawiedliwość, technologie sztucznej inteligencji mogą lepiej służyć wszystkim ludziom, niezależnie od ich pochodzenia czy tożsamości.



ZOBACZ: „Czy jesteśmy w stanie zapobiec dziedziczeniu naszych uprzedzeń przez sztuczną inteligencję?”

Podstawowe zasady etyki sztucznej inteligencji

Przejrzystość w systemach sztucznej inteligencji

Przejrzystość w sztucznej inteligencji odnosi się do zrozumienia mechanizmów podejmowania decyzji przez systemy AI. W miarę jak sztuczna inteligencja jest integrowana z takimi dziedzinami jak opieka zdrowotna i wymiar sprawiedliwości, przejrzystość staje się kluczowym elementem budowania zaufania, odpowiedzialności i uczciwości.

Dlaczego przejrzystość jest istotna

Przejrzystość umożliwia interesariuszom, takim jak deweloperzy, użytkownicy oraz organy regulacyjne, zrozumienie funkcjonowania systemów AI. Umożliwia użytkownikom kwestionowanie decyzji oraz buduje zaufanie. Jest to szczególnie istotne w obszarach wysokiego ryzyka, takich jak rekrutacja, udzielanie pożyczek oraz egzekwowanie prawa.

Kluczowe aspekty przejrzystości

- 1. Wyjaśnialność:** Umiejętność ludzi do pojmowania decyzji podejmowanych przez sztuczną inteligencję, co jest kluczowe w dziedzinach takich jak opieka zdrowotna i prawo.
- 2. Dostępność informacji:** Uczynienie algorytmów, źródeł danych i modeli przejrzystymi oraz dostępnymi do weryfikacji.
- 3. Audytowalność:** umożliwienie przeprowadzania audytów zewnętrznych w celu oceny systemów AI pod kątem uczciwości, stronniczości oraz ryzyka.

Wyzwania dotyczące przejrzystości

- **Złożoność algorytmów:** Wiele modeli sztucznej inteligencji (AI), takich jak głębokie uczenie, jest trudnych do całkowitego wyjaśnienia z uwagi na ich charakter „czarnej skrzynki”.
- **Modele zastrzeżone:** Niektóre przedsiębiorstwa utrzymują modele w tajemnicy z powodów komercyjnych, co ogranicza przejrzystość.
- **Ograniczenia techniczne:** Niektóre systemy mogą nie być całkowicie przejrzyste z powodu

ograniczeń technicznych.

Korzyści wynikające z przejrzystości

- 1. Zaufanie:** Przejrzysta sztuczna inteligencja kształtuje zaufanie, ujawniając proces podejmowania decyzji.
- 2. Odpowiedzialność:** Przejrzystość sprzyja pociąganiu jednostek do odpowiedzialności, gdy decyzje dotyczące sztucznej inteligencji okażą się niewłaściwe.
- 3. Uczciwość:** umożliwia identyfikację oraz korektę błędów, zapewniając rzetelność.

Strategie poprawy przejrzystości

- 1. Przejrzysta dokumentacja:** Zapewnienie zrozumiałych wyjaśnień dotyczących funkcjonowania modeli sztucznej inteligencji, w tym danych oraz procesów szkoleniowych.
- 2. Otwarte algorytmy:** udostępnianie kodu źródłowego lub szczegółów modelu do zewnętrznej analizy.
- 3. Oceny wpływu sztucznej inteligencji:** analiza wpływu społecznego, ekonomicznego i kulturowego sztucznej inteligencji oraz publikacja wyników.
- 4. Zaangażowanie interesariuszy:** angażowanie osób zainteresowanych w proces rozwoju sztucznej inteligencji, szczególnie w obszarach wrażliwych.

Wniosek

Przejrzystość stanowi fundament etycznego funkcjonowania sztucznej inteligencji. Umożliwia zrozumienie, buduje zaufanie oraz zapewnia odpowiedzialność. W miarę postępu w dziedzinie sztucznej inteligencji, priorytetowe traktowanie przejrzystości przyczyni się do zapewnienia, że technologie AI są sprawiedliwe, zrozumiałe i godne zaufania.



Podstawowe zasady etyki sztucznej inteligencji

Odpowiedzialność w zakresie sztucznej inteligencji

Odpowiedzialność w zakresie sztucznej inteligencji odnosi się do tego, kto jest odpowiedzialny, gdy systemy AI popełniają błędy lub wyrządzają szkody. W miarę jak AI coraz bardziej integruje się z codziennym życiem, precyzyjne określenie odpowiedzialności staje się kluczowe dla zapewnienia etycznego użytkowania oraz zapobiegania nadużyciom.

Dlaczego odpowiedzialność jest istotna

Systemy AI mają zdolność podejmowania decyzji, które wpływają na ludzkie życie, od aplikacji o pracę po leczenie. W przypadku awarii systemu AI, niezależnie od tego, czy wynika to z błędów projektowych, czy niewłaściwego użycia, odpowiedzialność zapewnia, że osoby odpowiedzialne za rozwój, wdrażanie i nadzorowanie technologii ponoszą konsekwencje jej skutków.

Kluczowe zagadnienia związane z odpowiedzialnością za sztuczną inteligencję

1. Kto ponosi odpowiedzialność?: W przypadku, gdy system sztucznej inteligencji wyrządzi szkodę, czy odpowiedzialność spoczywa na twórcach, użytkownikach, czy na firmie będącej właścicielem systemu? To zagadnienie jest kluczowe w obszarach takich jak np. autonomiczna jazda.

2. Przypisanie winy: W złożonych systemach AI powiązanie błędów z konkretną przyczyną może być wyzwaniem. Ustalenie, czy błąd wynika z wadliwego projektu, stronniczych danych czy niewłaściwego użytkowania, jest skomplikowane, co utrudnia określenie odpowiedzialności.

3. Współodpowiedzialność: W związku z tym, że w rozwój sztucznej inteligencji angażuje się wiele podmiotów (programistów, dostawców danych, przedsiębiorstw), ustalenie współodpowiedzialności jest kluczowe dla zapewnienia, że odpowiedzialność nie zostanie przerzucona na jedną stronę.

Wyzwania związane z odpowiedzialnością

- **Natura czarnej skrzynki:** Wiele modeli sztucznej inteligencji, szczególnie tych opartych na głębokim uczeniu, charakteryzuje się brakiem przejrzystości. Gdy decyzje są nie do wyjaśnienia, identyfikacja przyczyny błędu staje się trudna.

- **Zautomatyzowane decyzje:** W sytuacjach, gdy systemy AI podejmują decyzje autonomicznie, bez ludzkiej ingerencji, ustalenie odpowiedzialności staje się skomplikowane. Czy twórca powinien odpowiadać za działania AI, czy też sam system AI powinien ponosić część odpowiedzialności?

- **Luki w przepisach:** W wielu krajach brakuje regulacji dotyczących sztucznej inteligencji, co może utrudniać pociąganie osób fizycznych lub przedsiębiorstw do odpowiedzialności w przypadku nieprawidłowości.

Rozwiązania gwarantujące rozliczalność

1. Jasne zasady zarządzania: Ustanowienie wyraźnych ram definiujących odpowiedzialność na każdym etapie rozwoju i wdrażania sztucznej inteligencji może przyczynić się do zapewnienia rozliczalności.

2. Audyty AI: Regularne audyty zewnętrzne mogą zapewnić, że systemy AI funkcjonują w sposób etyczny i zgodny z zamierzonymi wytycznymi, co z kolei pomaga w identyfikacji potencjalnych błędów lub stronniczości.

3. Przejrzyste raportowanie: Twórcy oprogramowania oraz przedsiębiorstwa powinni ujawniać mechanizmy działania i procesy decyzyjne swoich systemów AI, aby zapewnić odpowiedzialność za osiągnięte wyniki.

4. Normy prawne i etyczne: Ustanowienie przejrzystych ram prawnych dotyczących zastosowania sztucznej inteligencji, szczególnie w sektorach o wysokim ryzyku, takich jak opieka zdrowotna, finanse i wymiar sprawiedliwości, może zapewnić, że osoby lub przedsiębiorstwa będą pociągnięte do odpowiedzialności w przypadku wystąpienia nieprawidłowości.

Odpowiedzialność za sztuczną inteligencję jest kluczowa dla jej etycznego zastosowania. Przez przypisanie odpowiedzialności, wdrożenie mechanizmów nadzoru oraz zapewnienie przejrzystości, możemy tworzyć systemy, które są zarówno efektywne, jak i sprawiedliwe. W miarę postępu w dziedzinie sztucznej inteligencji, zapewnienie odpowiedzialności będzie niezbędne dla utrzymania zaufania publicznego oraz ochrony przed potencjalnymi szkodami.

Podstawowe zasady etyki sztucznej inteligencji

Ochrona danych osobowych

W miarę jak integracja systemów sztucznej inteligencji z codziennym życiem staje się coraz bardziej zaawansowana, odpowiedzialne przetwarzanie danych osobowych nabiera kluczowego znaczenia. Od aplikacji monitorujących zdrowie po rozpoznawanie twarzy w przestrzeni publicznej, sztuczna inteligencja coraz częściej opiera się na danych wrażliwych. Dlatego istotne jest przestrzeganie standardów prywatności oraz ram prawnych, takich jak unijna ustawa o sztucznej inteligencji (AI Act) i ogólne rozporządzenie o ochronie danych (RODO). Ochrona tych danych jest niezbędna dla utrzymania zaufania użytkowników, zapobiegania nadużyciom oraz zgodności z europejskimi wartościami w zakresie etyki danych.

Dlaczego ochrona prywatności ma kluczowe znaczenie

Systemy sztucznej inteligencji (AI) często wymagają dostępu do danych osobowych, w tym informacji o lokalizacji, identyfikatorów biometrycznych, historii przeglądania oraz interakcji społecznych. Niewłaściwe wykorzystanie tych danych może prowadzić do dyskryminacji, kradzieży tożsamości lub nielegalnego profilowania. Zgodnie z ustawą o sztucznej inteligencji (AI Act), systemy AI uznawane za wysokiego ryzyka muszą spełniać rygorystyczne wymogi dotyczące jakości danych, ich zarządzania oraz przejrzystości. Naruszenia RODO, takie jak przetwarzanie danych bez podstawy prawnej lub brak zapewnienia praw osobom, których dane dotyczą, mogą skutkować poważnymi sankcjami. Twórcy i użytkownicy AI muszą zatem zapewnić, że dane osobowe są przetwarzane zgodnie z prawem, rzetelnie i w jasno określonym celu.

Kluczowe kwestie dotyczące ochrony danych osobowych

- 1. Gromadzenie danych i zgoda:** Użytkownicy muszą być wyraźnie poinformowani o tym, jakie dane osobowe są zbierane oraz w jakim celu. Zgoda musi być jednoznaczna i dobrowolna – zgoda dorozumiana lub łączona nie jest uznawana za ważną zgodnie z RODO.
- 2. Bezpieczeństwo danych:** Systemy sztucznej inteligencji powinny być wyposażone w solidne zabezpieczenia techniczne, takie jak szyfrowanie oraz kontrola dostępu. Każde naruszenie może wiązać się z obowiązkiem raportowania zgodnie z RODO w ciągu 72 godzin.
- 3. Anonimizacja danych:** Do trenowania modeli sztucznej inteligencji, które nie dotyczą osób, należy wykorzystywać wyłącznie dane zanonimizowane lub odpowiednio pseudonimizowane. Programiści muszą jednak unikać zestawów danych, które mogą być „ponownie identyfikowalne” i ujawniać tożsamość poszczególnych

osób poprzez wnioskowanie.

4. Wykorzystanie i własność danych: Przejrzystość w zakresie tego, kto jest właścicielem danych, kto je kontroluje oraz kto czerpie z nich korzyści, jest kluczowa. Zgodnie z RODO osoby fizyczne mają prawo do informacji na temat sposobu wykorzystywania ich danych – nawet w kontekście sztucznej inteligencji.

5. Przechowywanie i usuwanie danych: Dane osobowe nie mogą być przechowywane w nieskończoność. Systemy sztucznej inteligencji powinny zawierać mechanizmy usuwania danych, gdy nie są już potrzebne, oraz umożliwiać użytkownikom korzystanie z prawa do usunięcia danych (prawa do bycia zapomnianym).

Dane wrażliwe, takie jak pochodzenie rasowe lub etniczne, poglądy polityczne, przekonania religijne czy informacje dotyczące dzieci, nigdy nie powinny być wprowadzane do systemów sztucznej inteligencji, chyba że istnieje ku temu podstawa prawna oraz odpowiednie zabezpieczenia zgodnie z RODO. W większości placówek edukacyjnych i młodzieżowych należy bezwzględnie unikać wprowadzania takich danych.

Wytyczne etyczne w zakresie ochrony danych osobowych

- 1. Przejrzystość:** Użytkownicy powinni być zawsze informowani o tym, jakie dane są zbierane, w jaki sposób są wykorzystywane oraz czy stosowane jest profilowanie lub zautomatyzowane podejmowanie decyzji.
- 2. Minimalizacja danych:** Systemy sztucznej inteligencji powinny zbierać jedynie niezbędną minimalną ilość danych do ich funkcjonowania – umożliwia to ograniczenie ryzyka oraz zwiększenie zgodności z regulacjami.
- 3. Kontrola użytkownika:** Osoby fizyczne powinny mieć możliwość dostępu do swoich danych, ich modyfikacji lub usunięcia. Jest to fundamentalne prawo wynikające z RODO i kluczowe dla budowania etycznej wiedzy na temat sztucznej inteligencji.
- 4. Wyjaśnialność sztucznej inteligencji:** Wyjaśnialna sztuczna inteligencja powinna stać się normą, zwłaszcza gdy podejmowane decyzje mają wpływ na prawa lub możliwości jednostek. Użytkownicy mają prawo do istotnych informacji dotyczących logiki, która leży u podstaw zautomatyzowanych decyzji.
- 5. Ochrona danych w fazie projektowania:** Twórcy oprogramowania powinni integrować ochronę prywatności na każdym etapie projektowania systemu AI – od planowania po wdrożenie – zapewniając zgodność z RODO oraz nowymi regulacjami dotyczącymi AI.

Podstawowe zasady etyki sztucznej inteligencji

Etyka technologii deepfake

Deep fake'i to media generowane przez sztuczną inteligencję, takie jak obrazy, filmy czy pliki audio, które w niezwykle realistyczny sposób manipulują rzeczywistością. Te syntetyczne twory budzą poważne wątpliwości etyczne z uwagi na potencjalne nadużycia. Choć deep fake'i mogą być wykorzystywane w sposób kreatywny, ich zdolność do oszukiwania, szerzenia dezinformacji oraz naruszania prywatności czyni je kluczowym zagadnieniem w kontekście etyki sztucznej inteligencji.

Czym są deepfake'i?

Deep fake powstaje dzięki algorytmom głębokiego uczenia, które nakładają twarz jednej osoby na twarz innej lub modyfikują sposób mówienia, co sprawia, że wydaje się, iż ktoś wypowiada słowa, których nie powiedział. Realistyczna natura deep fake'ów uczyniła je potężnym narzędziem, zarówno w celach kreatywnych, jak i szkodliwych.

Obawy etyczne dotyczące deepfake'ów

1. Dezinformacja i manipulacja: Deepfake'i mogą być używane do manipulacji opinią publiczną lub szerzenia fałszywych informacji. Na przykład, były wykorzystywane do podszywania się pod postacie polityczne, co mogło wpływać na wyniki wyborów lub zmieniać postrzeganie społeczne.

2. Zgoda i prywatność: Tworzenie deep fake'ów bez czyjejkolwiek zgody narusza prywatność oraz autonomię jednostki. Wizerunki lub głosy osób mogą być wykorzystywane w sposób złośliwy, co może prowadzić do zniesławienia, nękania lub krzywdy emocjonalnej.

3. Odpowiedzialność: Anonimowość związana z tworzeniem deep fake'ów utrudnia ustalenie odpowiedzialności. Bez wyraźnego określenia odpowiedzialności trudno jest przeciwdziałać negatywnym skutkom deep fake'ów, takim jak szkody wizerunkowe czy stres emocjonalny.

4. Utrata zaufania: W miarę jak rośnie wyrafinowanie technologii deepfake, podważa to zaufanie do mediów cyfrowych. Gdy ludzie nie są w stanie odróżnić treści autentycznych od fałszywych, może to osłabić zaufanie do platform

internetowych, mediów oraz instytucji.

Wytyczne etyczne dotyczące obcowania z deepfake'ami

1. Przejrzystość: Systemy sztucznej inteligencji generujące deep fake'i powinny informować, kiedy treść jest sztuczna. Taka przejrzystość wspiera widzów w ocenie autentyczności prezentowanych materiałów.

2. Zgoda i własność: Osoby powinny mieć kontrolę nad tym, w jaki sposób ich wizerunek jest wykorzystywany w treściach generowanych przez sztuczną inteligencję. Należy uzyskać zgodę przed użyciem czyjegoś wizerunku lub głosu w deep fake'ach.

3. Wykrywanie i regulacje: Narzędzia detekcji oparte na sztucznej inteligencji mogą wspierać identyfikację deep fake'ów, a jednocześnie konieczne jest wprowadzenie regulacji, które będą zniechęcać do ich złośliwego wykorzystania. Pociągnięcie twórców do odpowiedzialności za szkodliwe treści jest kluczowe.

4. Edukacja i świadomość: Kształcenie społeczeństwa w zakresie deep fake'ów oraz promowanie umiejętności cyfrowych umożliwi ludziom krytyczną ocenę treści, z którymi stykają się w Internecie.

5. Etyczny rozwój sztucznej inteligencji: Twórcy powinni koncentrować się nie tylko na rozwijaniu technologii sztucznej inteligencji, ale również na jej etycznych implikacjach, zapewniając, że technologia deep fake jest stosowana w sposób odpowiedzialny.

Wniosek

Deep fake'i stwarzają skomplikowane wyzwania etyczne, ponieważ ich technologia może być zarówno twórcza, jak i szkodliwa. Aby zminimalizować ryzyko, kluczowe jest zapewnienie przejrzystości, zgody oraz odpowiedzialności w ich zastosowaniu. Dzięki odpowiednim regulacjom, edukacji i odpowiedzialnemu rozwojowi możemy zagwarantować, że treści generowane przez sztuczną inteligencję służą dobru publicznemu, a nie podważają prywatności i zaufania.

Wnioski: Kluczowe zasady etyki sztucznej inteligencji

W miarę jak sztuczna inteligencja odgrywa coraz bardziej znaczącą rolę w kształtowaniu naszego świata, zrozumienie oraz przestrzeganie podstawowych zasad etycznych związanych z jej rozwojem i zastosowaniem staje się kluczowe. Etyka w sztucznej inteligencji nie jest jedynie dodatkiem, lecz fundamentalnym elementem, który zapewnia, że sztuczna inteligencja służy ludzkości w sposób odpowiedzialny i korzystny.

Od zrozumienia stronniczości i dążenia do przejrzystości, poprzez zapewnienie rozliczalności, ochronę danych osobowych oraz rozważenie konsekwencji technologii takich jak deep fakes – każda zasada odgrywa kluczową rolę w kształtowaniu przyszłości sztucznej inteligencji. Te fundamentalne zasady kierują programistami, decydentami i użytkownikami w tworzeniu systemów sztucznej inteligencji, które są sprawiedliwe, odpowiedzialne i zgodne z wartościami społecznymi.

Ostatecznie, wspieranie kultury etycznej sztucznej inteligencji wymaga współpracy, edukacji oraz zaangażowania w nieustanną refleksję. W miarę wprowadzania innowacji i integracji sztucznej inteligencji z różnymi aspektami naszego życia, niezwykle istotne jest, aby pamiętać, że etyczne decyzje podejmowane dzisiaj wpłyną na przyszły wpływ sztucznej inteligencji na społeczeństwo. Przestrzegając tych fundamentalnych zasad, możemy zapewnić, że sztuczna inteligencja będzie wykorzystywana do promowania równości, przejrzystości, prywatności oraz odpowiedzialności, przynosząc korzyści wszystkim członkom społeczeństwa.



„Etyczne zastosowanie sztucznej inteligencji nie polega jedynie na realizacji tego, co jest technicznie możliwe, lecz na podejmowaniu działań, które są słuszne”.

Sundar Pichai,
Dyrektor generalny Alphabet (Google)



03

DYLEMATY ETYCZNE W DZIEDZINIE SZTUCZNEJ INTELIGENCJI



Dylematy etyczne związane ze sztuczną inteligencją

Ocena wyzwań etycznych związanych ze sztuczną inteligencją

Wraz z postępowaniem technologii sztucznej inteligencji (AI) pojawiają się coraz bardziej złożone wyzwania etyczne. Obejmują one nie tylko od dawna istniejące problemy, takie jak uprzedzenia i prywatność, ale także nowsze, takie jak halucynacje generowane przez sztuczną inteligencję generatywną (np. ChatGPT lub Google Gemini), nadmierne poleganie na błędnych danych wejściowych oraz brak świadomości kontekstowej w systemach decyzyjnych. Na przykład model językowy może z dużą pewnością generować błędne porady prawne lub medyczne, a także niewłaściwie interpretować wskazówki kulturowe w zadaniach moderacji treści. Ocena takich zagrożeń jest kluczowa dla zapewnienia odpowiedzialnego, bezpiecznego i sprawiedliwego korzystania ze sztucznej inteligencji. Obecnie wiele narzędzi i frameworków wspiera programistów oraz edukatorów w poruszaniu się po kwestiach etycznych związanych z systemami AI.

Zaufana sztuczna inteligencja

Wytyczne Komisji Europejskiej dotyczące etyki w zakresie zaufanej sztucznej inteligencji definiują kluczowe zasady, które systemy sztucznej inteligencji muszą spełniać, aby mogły być uznawane za zgodne z prawem, etyczne i solidne:

- **Przejrzystość:** Systemy sztucznej inteligencji powinny być zrozumiałe dla użytkowników. Na przykład, sztuczna inteligencja do gier „Cicero” firmy Meta została doceniona za integrację przejrzystości z strategią języka naturalnego – jednocześnie jednak zwrócono uwagę na ryzyko manipulacji w systemach sztucznej inteligencji opartych na języku, jeśli nie są one wystarczająco przejrzyste.
- **Odpowiedzialność:** Programiści są zobowiązani do ponoszenia odpowiedzialności za decyzje oraz wyniki sztucznej inteligencji. Narzędzia takie jak Model Cards (Google) oraz FactSheets (IBM) dostarczają dokumentacji, która wyjaśnia cele i ograniczenia systemów sztucznej inteligencji.
- **Prywatność i ochrona danych:** Narzędzia sztucznej inteligencji, takie jak rozpoznawanie twarzy, muszą być zgodne z RODO oraz przepisami krajowymi. Przypadek Clearview AI, w którym dane biometryczne zostały zebrane bez zgody, ilustruje, jak sztuczna inteligencja może naruszać prawo do prywatności w obliczu niewystarczających zabezpieczeń.
- **Niedyskryminacja i uczciwość:** Sztuczna inteligencja stosowana w narzędziach rekrutacyjnych (np. porzucony algorytm rekrutacyjny firmy Amazon) ujawniła, w jaki sposób szkolenie oparte na stronicznych danych historycznych może potęgować dyskryminację ze względu na płeć lub rasę.

- **Bezpieczeństwo:** Systemy autonomiczne, takie jak wersja beta systemu Full Self-Driving firmy Tesla, budzą obawy dotyczące bezpieczeństwa, gdy podejmowanie decyzji zawodzi w obliczu rzeczywistej złożoności. Luki w zabezpieczeniach asystentów głosowych, takich jak Alexa, również były wykorzystywane do uzyskiwania nieautoryzowanego dostępu.

Narzędzia i ramy etyki sztucznej inteligencji

Organizacje oraz badacze opracowali szereg praktycznych narzędzi, które wspierają zarządzanie ryzykiem związanym ze sztuczną inteligencją:

1. **Oceny skutków:** Oceny skutków sztucznej inteligencji – takie jak proponowana przez UE klasyfikacja ryzyka związanego ze sztuczną inteligencją na mocy ustawy o sztucznej inteligencji – analizują zagrożenia społeczne przed ich wdrożeniem. Na przykład, w niektórych europejskich miastach zakazano stosowania narzędzi predykcyjnych w policji po tym, jak analizy ujawniły nieproporcjonalne szkody dla mniejszości.
2. **Narzędzia do wykrywania uprzedzeń:** Narzędzia takie jak AI Fairness 360 firmy IBM, Fairlearn firmy Microsoft oraz „What-If Tool” firmy Google umożliwiają programistom ocenę modeli sztucznej inteligencji pod kątem uprzedzeń. Na przykład, algorytm COMPAS stosowany w amerykańskich sądach wykazuje różnice w prognozowaniu recydywy w zależności od rasy, co podkreśla konieczność przeprowadzania audytów w celu identyfikacji uprzedzeń przed ich wdrożeniem w praktyce.
3. **Ramy zarządzania sztuczną inteligencją:** Zasady OECD dotyczące sztucznej inteligencji (AI) oraz Rekomendacja UNESCO w sprawie etyki sztucznej inteligencji (AI) dostarczają niewiążących, lecz wpływowych wytycznych dotyczących wiarygodnej, zorientowanej na człowieka sztucznej inteligencji. Ramy te wspierają organizacje w tworzeniu wewnętrznych polityk – na przykład wymagających nadzoru ludzkiego w zautomatyzowanych systemach podejmowania decyzji stosowanych w edukacji lub usługach publicznych.
4. **Wyjaśnialność i interpretowalność:** Techniki wyjaśnialnej sztucznej inteligencji (XAI), takie jak LIME i SHAP, mogą ujawnić, które cechy wejściowe miały największy wpływ na podejmowane decyzje. Jest to kluczowe w kontekście sztucznej inteligencji w opiece zdrowotnej, gdzie modele typu „czarna skrzynka”, sugerujące metody leczenia bez odpowiedniego uzasadnienia (np. IBM Watson for Oncology), spotykają się z krytyką za brak przejrzystości oraz niską dokładność w rzeczywistych warunkach.

Dylematy etyczne związane ze sztuczną inteligencją

Ocena wyzwań etycznych związanych ze sztuczną inteligencją

Zasady przewodnie etycznego projektowania sztucznej inteligencji.

Aby skuteczniej ocenić wyzwania etyczne, twórcy oprogramowania oraz praktycy sztucznej inteligencji mogą opierać się na kilku zasadach przewodnich, które wspierają proces podejmowania decyzji etycznych:

- **Projektowanie zorientowane na człowieka:** Systemy AI powinny być tworzone z uwzględnieniem dobra człowieka. Oznacza to, że bezpieczeństwo użytkownika, uczciwość oraz poszanowanie praw człowieka muszą być priorytetem w każdym etapie rozwoju AI.
- **Rozwój wspierający włączenie społeczne:** Etyczne projektowanie sztucznej inteligencji wymaga uwzględnienia różnorodnych perspektyw, szczególnie tych grup, które są marginalizowane lub niedostatecznie reprezentowane, aby zapewnić, że systemy sztucznej inteligencji będą służyć wszystkim w równy sposób.
- **Ciągły monitoring:** Kwestie etyczne powinny być traktowane nie jako jednorazowe zagadnienie, lecz jako proces nieprzerwany. Systematyczne monitorowanie i audyt systemów AI są kluczowe dla zapewnienia ich etyczności przez cały cykl życia.

Ocena wyzwań etycznych w obszarze sztucznej inteligencji (AI) wymaga integracji narzędzi technicznych, ram oraz zasad filozoficznych. Te narzędzia umożliwiają programistom, decydentom i innym zainteresowanym stronom analizę ryzyka oraz korzyści związanych z AI, zapewniając, że jej zastosowanie jest zgodne z normami etycznymi i wartościami społecznymi. W miarę postępu AI, te ramy będą kluczowe dla kierowania jej odpowiedzialnym rozwojem i zastosowaniem.

PRZECZYTAJ: Etyczne zagadnienia sztucznej inteligencji (AI) w szkolnictwie wyższym



Dylematy etyczne związane ze sztuczną inteligencją

Sztuczna inteligencja a przyszłość społeczeństwa

Sztuczna inteligencja przestała być futurystyczną koncepcją; zmienia sposób, w jaki żyjemy, pracujemy i komunikujemy się. Wraz z postępem technologii AI, niesie ona ze sobą zarówno ekscytujące możliwości, jak i złożone dylematy, które mogą mieć istotny wpływ na społeczeństwo. Wyzwania te są szczególnie istotne, ponieważ AI zaczyna oddziaływać na kluczowe decyzje dotyczące zatrudnienia, prywatności, a nawet zarządzania. Niniejszy podrozdział analizuje niektóre z najpoważniejszych dylematów etycznych związanych z AI, w tym kwestie dotyczące przenoszenia miejsc pracy, nadzoru oraz podejmowania decyzji przy użyciu AI.

Przemieszczanie pracowników oraz przyszłość pracy

Jednym z najczęściej dyskutowanych dylematów etycznych związanych ze sztuczną inteligencją (AI) jest jej potencjał do eliminacji miejsc pracy. Technologie AI, takie jak automatyzacja, robotyka i uczenie maszynowe, stają się coraz bardziej zdolne do realizacji zadań tradycyjnie wykonywanych przez ludzi. Dotyczy to zatrudnienia w przemyśle, transporcie, obsłudze klienta, a nawet w takich sektorach jak opieka zdrowotna i prawo. Choć sztuczna inteligencja może przyczynić się do wzrostu efektywności i produktywności, rodzi również obawy, czy maszyny zastąpią ludzi, co może prowadzić do powszechnej utraty miejsc pracy.

Obawy dotyczące utraty miejsc pracy w wyniku rozwoju sztucznej inteligencji są uzasadnione. Na przykład, rozwój pojazdów autonomicznych stwarza ryzyko utraty zatrudnienia dla kierowców, a automatyzacja w magazynach i fabrykach może zastąpić pracę fizyczną. Zgodnie z badaniem McKinsey, do 2030 roku nawet 800 milionów pracowników na całym świecie może stracić pracę z powodu automatyzacji i technologii sztucznej inteligencji.

Jednak sztuczna inteligencja stwarza również możliwości tworzenia miejsc pracy w rozwijających się sektorach, takich jak etyka sztucznej inteligencji, inżynieria uczenia maszynowego oraz analiza danych. Wyzwaniem będzie zapewnienie pracownikom umiejętności niezbędnych do odniesienia sukcesu w świecie, w którym sztuczna

inteligencja odgrywa kluczową rolę. W miarę postępu sztucznej inteligencji, programy edukacyjne i przekwalifikowujące będą miały fundamentalne znaczenie dla przygotowania pracowników do nowych ról, minimalizując negatywny wpływ utraty zatrudnienia.

Nadzór oraz prywatność

Kolejnym istotnym dylematem etycznym związanym ze sztuczną inteligencją jest problem nadzoru. Technologie sztucznej inteligencji, takie jak oprogramowanie do rozpoznawania twarzy, eksploracja danych oraz analityka predykcyjna, mają potencjał do stworzenia środowiska nieustannego monitoringu. Rządy i korporacje coraz częściej wykorzystują sztuczną inteligencję do śledzenia zachowań, ruchów, a nawet emocji jednostek, często bez ich wiedzy i zgody.

Z jednej strony, nadzór oparty na sztucznej inteligencji może zwiększyć bezpieczeństwo, wspierać identyfikację przestępców, a nawet zapobiegać potencjalnym zagrożeniom. Rodzi jednak poważne obawy dotyczące prywatności oraz możliwości nadużyć. Masowy nadzór obywateli przez rządy może prowadzić do naruszeń wolności osobistych i praw człowieka. Technologie sztucznej inteligencji są również coraz częściej wykorzystywane w sektorze prywatnym, gdzie przedsiębiorstwa monitorują zachowania konsumentów w celach marketingowych, co prowadzi do bardziej inwazyjnej formy nadzoru, mogącej naruszać prywatność.

Dylemat etyczny polega na osiągnięciu równowagi między korzyściami wynikającymi z nadzoru opartego na sztucznej inteligencji a ochroną prywatności jednostek. Istotna jest potrzeba jasnych regulacji oraz transparentnych zasad dotyczących zastosowania sztucznej inteligencji w nadzorze. W niektórych krajach wprowadzono przepisy, takie jak Ogólne Rozporządzenie o Ochronie Danych (RODO) w Unii Europejskiej, mające na celu ograniczenie zakresu nadzoru oraz ochronę prawa do prywatności. Niemniej jednak, konieczne są dalsze globalne działania, aby zapewnić, że praktyki nadzoru z wykorzystaniem sztucznej inteligencji nie naruszają podstawowych wolności.

Dylematy etyczne związane ze sztuczną inteligencją

Sztuczna inteligencja a przyszłość społeczeństwa

Sztuczna inteligencja w procesie podejmowania decyzji

Rosnąca rola sztucznej inteligencji w procesach decyzyjnych stwarza poważny dylemat etyczny. Systemy sztucznej inteligencji są coraz częściej stosowane w takich dziedzinach jak rekrutacja, egzekwowanie prawa, opieka zdrowotna, a nawet wymiar sprawiedliwości. Te systemy analizują ogromne ilości danych i wykorzystują algorytmy do podejmowania decyzji, które wcześniej były w gestii ludzi. Na przykład, sztuczna inteligencja jest używana do oceny predyspozycji kandydatów do pracy, przewidywania prawdopodobieństwa recydywy, a także diagnozowania schorzeń.

Chociaż sztuczna inteligencja jest zdolna do przetwarzania informacji z prędkością i na skalę, które przewyższają możliwości ludzkie, poleganie na algorytmach w podejmowaniu decyzji rodzi poważne wątpliwości etyczne. Kluczowym problemem jest potencjalna stronniczość w procesie decyzyjnym sztucznej inteligencji. Jeśli systemy sztucznej inteligencji są szkolone na danych obciążonych stronniczością, mogą utrzymywać, a nawet wzmacniać istniejące nierówności społeczne. Na przykład, system sztucznej inteligencji stosowany w procesie rekrutacji może faworyzować kandydatów z określonych grup demograficznych, jednocześnie ignorując inne, co prowadzi do dyskryminacji.

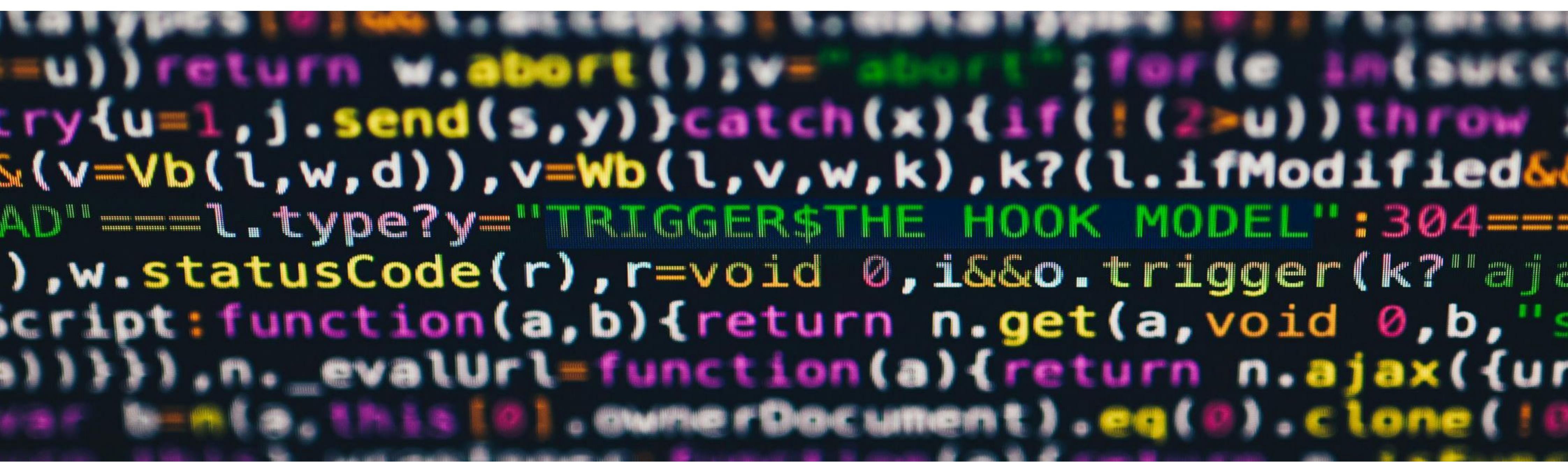
Kwestia odpowiedzialności odgrywa kluczową rolę w procesie podejmowania decyzji w dziedzinie sztucznej inteligencji. Gdy system AI podejmuje błędną lub szkodliwą decyzję – na przykład niesłuszne zwolnienie pracownika lub skazanie niewinnej osoby – kto powinien ponosić odpowiedzialność? Czy odpowiedzialność powinni

ponosić twórcy, organizacje korzystające z AI, czy sama sztuczna inteligencja? Obecnie przepisy i regulacje dotyczące odpowiedzialności w kontekście AI są wciąż niedostatecznie rozwinięte, co wskazuje na potrzebę wprowadzenia bardziej przejrzystych ram prawnych w celu określenia odpowiedzialności w takich sytuacjach.

Wpływ sztucznej inteligencji na zarządzanie.

Wreszcie, potencjalny wpływ sztucznej inteligencji na zarządzanie stwarza nowy dylemat etyczny. Dzięki zdolności sztucznej inteligencji do analizy i przewidywania trendów na podstawie ogromnych zbiorów danych, istnieje możliwość wpływania na decyzje polityczne i strategiczne. Sztuczna inteligencja mogłaby być wykorzystywana do optymalizacji działań administracji publicznej, poprawy usług publicznych oraz tworzenia inteligentniejszych miast. Niemniej jednak, rosnące zastosowanie sztucznej inteligencji w zarządzaniu może również prowadzić do nierównowagi sił oraz niezamierzonych konsekwencji.

Na przykład systemy sztucznej inteligencji mogą być wykorzystywane do manipulowania opinią publiczną poprzez ukierunkowane reklamy polityczne lub fałszywe wiadomości, co podważa procesy demokratyczne. Ponadto, stosowanie systemów sztucznej inteligencji do podejmowania decyzji politycznych bez odpowiedniego nadzoru i przejrzystości może prowadzić do utraty zaufania społecznego do systemu. Dylemat etyczny w tym przypadku polega na zapewnieniu, że sztuczna inteligencja jest wykorzystywana w zarządzaniu w sposób promujący przejrzystość, uczciwość i odpowiedzialność, a jednocześnie chroniący prawa i wolności jednostki.



Wnioski: dylematy etyczne w obszarze sztucznej inteligencji

Jak już omówiono w tym rozdziale, wyzwania etyczne związane ze sztuczną inteligencją są znaczące i złożone. Od redukcji miejsc pracy i obaw dotyczących nadzoru, po transparentność i odpowiedzialność systemów sztucznej inteligencji, dylematy te dotyczą niemal każdego aspektu współczesnego życia. Potencjał sztucznej inteligencji do przekształcania społeczeństwa stwarza ogromne możliwości, ale wymaga również starannego rozważenia konsekwencji jej zastosowania.

Sprostanie tym wyzwaniom etycznym wymaga nieustannego dialogu, współpracy oraz proaktywnych strategii. Potencjał sztucznej inteligencji (AI) do rewolucjonizowania branż, naruszania prywatności i wpływania na zarządzanie wymaga ustanowienia ram odpowiedzialności, uczciwości i przejrzystości. W miarę dalszej integracji AI z kluczowymi procesami decyzyjnymi, istotne jest uwzględnienie szerszych implikacji społecznych, aby zapewnić, że technologia ta jest zgodna z naszymi wartościami i służy dobru wspólnemu.

Ostatecznie przyszłość sztucznej inteligencji (AI) nie sprowadza się jedynie do postępu technologicznego; obejmuje również sposób, w jaki my, jako społeczeństwo, zdecydujemy się rozwiązać te dylematy etyczne. Wyposażając pracowników młodzieżowych, edukatorów oraz przyszłe pokolenia w narzędzia i wiedzę do krytycznej oceny wpływu sztucznej inteligencji, możemy kształtować przyszłość, w której AI będzie wykorzystywana w sposób odpowiedzialny i etyczny, zapewniając, że jej korzyści dotrą do wszystkich sektorów społeczeństwa, nie naruszając naszych podstawowych wartości.



„Prawdziwe pytanie nie brzmi, czy maszyny myślą, lecz czy myślą ludzie.”

**B.F. Skinner,
znany psycholog oraz
behawiorysta.**



04

PRZEPROWADZENIE MŁODZIEŻY PRZEZ DYSKUSJE NA TEMAT ETYKI SZTUCZNEJ



Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

Ułatwienie dyskusji etycznych

Zorganizowanie dyskusji na temat etyki sztucznej inteligencji może być wyzwaniem, szczególnie w kontekście złożonych zagadnień, takich jak uprzedzenia, prywatność i odpowiedzialność. Niemniej jednak, zaangażowanie młodych ludzi w te rozmowy jest kluczowe dla rozwijania ich świadomości oraz krytycznego myślenia. Oto kilka strategii, które mogą wspierać pracowników młodzieżowych i edukatorów w prowadzeniu wartościowych i konstruktywnych dyskusji na temat etycznych dylematów związanych z AI:

1. Tworzenie bezpiecznego i dostępnego środowiska

Aby młodzież mogła angażować się w kwestie etyczne związane ze sztuczną inteligencją, niezbędna jest przestrzeń, w której będzie mogła swobodnie wyrażać swoje opinie, zadawać pytania i kwestionować idee. Stworzenie otwartej i pełnej szacunku atmosfery sprzyja zapewnieniu wszystkim uczestnikom poczucia bezpieczeństwa oraz możliwości dzielenia się swoimi przemyśleniami bez obawy przed oceną. Zachęcaj do aktywnego słuchania i empatii, aby wspierać dialog oparty na szacunku.

2. Powiązanie etyki sztucznej inteligencji z rzeczywistymi przykładami.

Wykorzystanie rzeczywistych przykładów zastosowań sztucznej inteligencji może pomóc młodym ludziom powiązać abstrakcyjne koncepcje etyczne z ich codziennym życiem. Należy omówić powszechne zastosowania sztucznej inteligencji, takie jak algorytmy mediów społecznościowych, technologia rozpoznawania twarzy oraz narzędzia rekrutacyjne oparte na sztucznej inteligencji. Ujmując dylematy etyczne w kontekście osobistych doświadczeń, młodzież może lepiej zrozumieć wpływ sztucznej inteligencji na ich życie społeczne i zawodowe. Na przykład, wyjaśnienie, w jaki sposób algorytmy sztucznej inteligencji kształtują treści w ich kanałach, może zainicjować dyskusję na temat stronniczości algorytmów, obaw o prywatność oraz odpowiedzialności firm technologicznych.

3. Zachęcanie do empatii poprzez odgrywanie ról oraz scenariusze.

Aby pogłębić zrozumienie, wykorzystaj scenariusze fabularne, w których młodzież może przyjąć różne perspektywy. Na przykład, niech wcielą się w rolę programistów, użytkowników lub decydentów AI, którzy podejmują decyzje dotyczące rozwoju lub użytkowania systemu SI. Ta metoda sprzyja empatii, ponieważ pozwala uczestnikom rozważyć dylematy etyczne z różnych punktów widzenia oraz zastanowić się, jak ich decyzje mogą wpłynąć na innych. Pomaga im również zrozumieć, że etyka AI nie dotyczy wyłącznie technologii, lecz także ludzkiego życia i dobrobytu społeczeństwa.

4. Promowanie pytań otwartych.

Zadawaj pytania otwarte, które skłaniają do refleksji i pobudzają głębsze myślenie. Zamiast formułować pytania, na które można odpowiedzieć jedynie „tak” lub „nie”, zachęcaj młodych ludzi do analizy własnych przekonań oraz rozważania szerszych konsekwencji decyzji podejmowanych w kontekście sztucznej inteligencji. Na przykład:

- „Co powinno się wydarzyć, jeśli system sztucznej inteligencji podejmie decyzję, która wyrządzi komuś szkodę? Kto ponosi odpowiedzialność?”
- „Czy sądzisz, że sztuczna inteligencja powinna być stosowana do monitorowania zachowań ludzi w przestrzeni publicznej? Dlaczego tak, a dlaczego nie?”
- „W jaki sposób możemy zapewnić, że systemy sztucznej inteligencji będą sprawiedliwe i nie będą dyskryminować wybranych grup ludzi?”

Odpowiedzi na tego rodzaju pytania wymagają od młodych ludzi krytycznego myślenia, analizy różnych czynników oraz omówienia złożoności etycznych aspektów sztucznej inteligencji.



Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

Organizowanie dyskusji etycznych

5. Uznawanie różnych perspektyw

Ważne jest, aby zaznajomić młodzież z różnorodnymi perspektywami, zwłaszcza że wpływ sztucznej inteligencji obejmuje różne sektory i kultury. Prezentuj różne punkty widzenia, czy to poprzez prelegentów, artykuły, czy filmy, od ekspertów ds. sztucznej inteligencji, etyków, aktywistów oraz technologów z różnych środowisk. To pomoże młodym ludziom dostrzec wielowymiarowy charakter wyzwań etycznych związanych ze sztuczną inteligencją oraz rozważyć szersze implikacje społeczne.

6. Objasnienie koncepcji wiarygodnej sztucznej inteligencji

Wprowadzając wyzwania etyczne, istotne jest również wprowadzenie koncepcji „godnej zaufania sztucznej inteligencji”, która oznacza, że systemy AI są przejrzyste, odpowiedzialne i projektowane z uwzględnieniem aspektów etycznych. Należy pomóc młodym ludziom zrozumieć, jak sztuczna inteligencja może być rozwijana i wykorzystywana w sposób odpowiedzialny, a także podkreślić znaczenie zaufania do technologii. Wykorzystaj przykłady z dziedziny etyki sztucznej inteligencji, takie jak zasady Google dotyczące sztucznej inteligencji czy wytyczne Komisji Europejskiej dotyczące godnej zaufania sztucznej inteligencji, aby pokierować dyskusją na temat tego, jak systemy AI mogą być jednocześnie innowacyjne i etyczne.

7. Zachęcanie do dyskusji inicjowanych przez młodzież

Daj młodym ludziom możliwość prowadzenia dyskusji oraz formułowania własnych pytań dotyczących etycznych implikacji sztucznej inteligencji. Zapewnij im ramy do podejmowania decyzji etycznych, takie jak zasady uczciwości, prywatności i przejrzystości, oraz zachęcaj ich do stosowania tych ram w swoich dyskusjach. Ta autonomia w podejmowaniu decyzji sprzyja krytycznemu myśleniu i wzmacnia przekonanie, że nie są jedynie biernymi odbiorcami informacji, lecz aktywnymi uczestnikami kształtowania przyszłości sztucznej inteligencji.



Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

Zachęta do krytycznego myślenia

Zachęcanie do krytycznego myślenia stanowi jedno z najskuteczniejszych narzędzi, z których mogą korzystać edukatorzy oraz pracownicy młodzieżowi, aby angażować młodych ludzi w dyskusje dotyczące etyki sztucznej inteligencji. Krytyczne myślenie umożliwia jednostkom ocenę etycznych implikacji technologii sztucznej inteligencji, kwestionowanie założeń oraz refleksję nad długofalowymi konsekwencjami zastosowań sztucznej inteligencji w społeczeństwie. W tym podrozdziale przedstawimy strategie rozwijania krytycznego myślenia wśród młodzieży oraz zachęcania ich do kwestionowania etycznego wymiaru sztucznej inteligencji.

1. Wspieranie postawy krytycznej

Jednym z kluczowych aspektów krytycznego myślenia jest zdolność do formułowania pytań, które kwestionują status quo i zgłębiają istotne kwestie etyczne. Zachęcaj młodych ludzi do przyjęcia postawy krytycznej, skłaniając ich do refleksji nad tym, w jaki sposób i dlaczego sztuczna inteligencja wpływa na społeczeństwo. Zamiast po prostu akceptować neutralność systemów AI lub ich funkcjonowanie zgodnie z zamierzeniem, należy nauczyć młodych ludzi zadawania pytań:

- „Kto odniesie korzyści z tego systemu sztucznej inteligencji, a kto może ponieść straty?”
- „Jakie wartości są obecne w tej technologii sztucznej inteligencji?”
- „W jaki sposób możemy mieć pewność, że ta technologia respektuje prywatność i uczciwość?”
- „Jakie są niezamierzone skutki takiego zastosowania sztucznej inteligencji?”

Tego rodzaju dociekliwe pytania stymulują głębsze zrozumienie społecznej roli sztucznej inteligencji oraz skłaniają do refleksji nad kwestiami etycznymi, takimi jak stronnictwo, uczciwość i potencjalne nadużycia.

2. Wykorzystanie studiów przypadków do podważania założeń

Studia przypadków z życia wzięte stanowią znakomity sposób na stymulowanie krytycznego myślenia oraz

angażowanie młodych ludzi w analizę etycznych implikacji sztucznej inteligencji. Przedstaw przypadki, w których systemy sztucznej inteligencji wywołały niezamierzone lub kontrowersyjne skutki, takie jak algorytmiczna stronnictwość, inwigilacja czy rozprzestrzenianie dezinformacji. Na przykład:

- **Sztuczna inteligencja w procesie rekrutacji:** Przeanalizuj przypadek systemów rekrutacyjnych opartych na sztucznej inteligencji, które niezamierzenie preferują jedną płć lub grupę etniczną, co rodzi wątpliwości dotyczące ich sprawiedliwości i obiektywności.
- **Deepfake’i w mediach:** Zademonstrowano przypadki, w których technologia deepfake została użyta do generowania wprowadzających w błąd lub szkodliwych treści, co wywołało obawy o zaufanie i autentyczność w erze cyfrowej.
- **Systemy monitoringu:** Analiza zagadnień etycznych dotyczących narzędzi monitorujących opartych na sztucznej inteligencji, stosowanych w przestrzeni publicznej, a także równowagi pomiędzy bezpieczeństwem a prywatnością.

Analizując te przypadki, młodzi ludzie są zachęceni do kwestionowania nie tylko samej technologii, ale również założeń i motywacji, które ją napędzają. Studia przypadków skłaniają do krytycznej oceny, w jaki sposób systemy sztucznej inteligencji mogą utrzymywać istniejące nierówności lub generować nowe dylematy etyczne.



ZOBACZ: „Etyka sztucznej inteligencji oraz zaangażowanie dzieci i młodzieży”

Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

Interaktywne działania angażujące uczestników

Interaktywne działania stanowią kluczowe narzędzia do angażowania młodzieży w dyskusje dotyczące etyki sztucznej inteligencji. Łącząc teorię z praktycznym doświadczeniem, te inicjatywy umożliwiają młodym ludziom zgłębianie etycznych aspektów sztucznej inteligencji w sposób bardziej dynamiczny i angażujący. Praktyczne ćwiczenia oraz scenariusze fabularne mogą wspierać młodzież w krytycznej analizie rzeczywistych dylematów etycznych, zrozumieniu różnorodnych perspektyw oraz rozwijaniu umiejętności prowadzenia przemyślanych, merytorycznych debat na temat implikacji technologii sztucznej inteligencji.

W tym podrozdziale zaprezentujemy szereg interaktywnych działań, które mają na celu stymulowanie myślenia, prowokowanie dyskusji oraz wspieranie młodych ludzi w aktywnym zaangażowaniu się w kwestie etyczne związane ze sztuczną inteligencją.

1. Debata na temat etyki sztucznej inteligencji: znaczenie sztucznej inteligencji w społeczeństwie

Jednym z najefektywniejszych sposobów angażowania młodzieży są zorganizowane debaty. Tego rodzaju aktywność umożliwia im rozważenie odmiennych perspektyw oraz obronę swoich poglądów, co sprzyja krytycznemu myśleniu o kwestiach etycznych związanych ze sztuczną inteligencją.

Cel: analiza różnych dylematów etycznych związanych ze sztuczną inteligencją oraz zachęcenie uczestników do refleksji nad korzyściami i wyzwaniami, które niesie ze sobą technologia sztucznej inteligencji.

Instrukcje:

- Podziel grupę na dwa zespoły: jeden zespół będzie bronił technologii AI, podczas gdy drugi będzie się im sprzeciwiał, koncentrując się na zagadnieniach etycznych, takich jak prywatność, stronniczość oraz potencjalna utrata miejsc pracy.
- Zbadaj temat: Zapewnij każdemu zespołowi odpowiednie materiały i zasoby dotyczące dylematów etycznych związanych ze sztuczną inteligencją, takich jak naruszenia prywatności, stronniczość algorytmiczna oraz wpływ na zatrudnienie.
- Ustal zasady debaty: Każdy zespół otrzyma określony czas na przedstawienie swoich argumentów, po którym nastąpi sesja repliki oraz końcowe podsumowanie.
- Pytania pomocnicze: Skorzystaj z pytań takich jak „Czy sztuczną inteligencję można etycznie włączyć do systemów publicznych?” lub „Jakie są potencjalne zagrożenia związane ze sztuczną inteligencją w procesie podejmowania decyzji?”, aby ukierunkować debatę.

Rezultat: Debata skłoni młodych ludzi do krytycznego analizy etycznych implikacji sztucznej inteligencji oraz wzmocni ich zdolność do formułowania własnych poglądów. Poprzez konfrontację z różnorodnymi punktami widzenia, nauczą się oceniać złożone kwestie etyczne z wielu perspektyw.

Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

2. Sztuczna inteligencja w rzeczywistości: analiza przypadku

Studia przypadków oferują cenne spostrzeżenia dotyczące praktycznych zastosowań sztucznej inteligencji (AI) oraz jej etycznych implikacji. Analizując rzeczywiste przykłady wykorzystania technologii AI, młodzi ludzie mogą lepiej zrozumieć, w jaki sposób AI może zarówno przynosić korzyści społeczeństwu, jak i rodzić dylematy etyczne.

Cel: Umożliwienie uczestnikom zapoznania się z rzeczywistymi przykładami zastosowań sztucznej inteligencji oraz omówienie ich konsekwencji etycznych.

Instrukcje:

- **Przedstaw studium przypadku:** Wybierz przypadek, który uwypukla dylemat etyczny w obszarze sztucznej inteligencji. Na przykład, możesz omówić przypadek systemu rozpoznawania twarzy stosowanego przez organy ścigania lub narzędzie rekrutacyjne oparte na sztucznej inteligencji, które nieumyślnie dyskryminuje określone grupy kandydatów.
- **Dyskusja grupowa:** Po zaprezentowaniu studium przypadku podziel grupę na mniejsze zespoły i poproś je o omówienie poruszonych kwestii etycznych. Zachęć uczestników do refleksji nad następującymi pytaniami:
 - Jakie kwestie etyczne są związane z tą technologią sztucznej inteligencji?
 - Kogo dotyczą te zagadnienia (np. jednostki, społeczności, społeczeństwo)?
 - Jakie działania można podjąć w celu rozwiązania tych problemów?
- **Przedstawienie ustaleń:** Każda grupa powinna zaprezentować swoją analizę, koncentrując się na proponowanych rozwiązaniach etycznych dylematów przedstawionych w studium przypadku.

Rezultat: Ta aktywność umożliwi młodym ludziom rozwijanie umiejętności analitycznych poprzez ocenę technologii AI z perspektywy etycznej. Zachęci ich także do refleksji nad tym, w jaki sposób mogliby praktycznie i wykonalnie rozwiązać wyzwania etyczne związane z AI.



CZYTAJ: Sztuczna inteligencja - przykłady dylematów etycznych.

Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

3. Odgrywanie ról: Kto ponosi odpowiedzialność?

Odgrywanie ról może wspierać młodych ludzi w zrozumieniu perspektyw różnych interesariuszy zaangażowanych w dylematy etyczne związane ze sztuczną inteligencją. To ćwiczenie umożliwia uczestnikom refleksję nad tym, jak różne osoby lub organizacje mogą podchodzić do kwestii etycznych oraz jaki wpływ ich decyzje mogą wywierać na innych.

Cel: Angażowanie młodzieży w badanie kwestii odpowiedzialności i rozliczalności w systemach sztucznej inteligencji.

Instrukcje:

- Przypisz role: Przydziel uczestnikom różnorodne role do odegrania w scenariuszu, w którym system sztucznej inteligencji generuje dylemat etyczny. Przykłady ról obejmują:
 - Programista, który opracował system sztucznej inteligencji.
 - Organ regulacyjny odpowiedzialny za nadzorowanie zastosowania sztucznej inteligencji.
 - Konsument, którego decyzje są kształtowane przez wybory dokonywane przez sztuczną inteligencję.
 - Aktywista promujący odpowiedzialne wykorzystanie sztucznej inteligencji.
- Przygotowanie scenariusza: Przedstaw scenariusz, na przykład stronniczy system sztucznej inteligencji (AI) wykorzystywany do zatwierdzania pożyczek, który nieproporcjonalnie często odmawia przyznania kredytu osobom z określonych społeczności. Uczestnicy będą dyskutować na temat odpowiedzialności za szkody wyrządzone przez system AI.
- Dyskusja i podejmowanie decyzji: Każdy z uczestników powinien zaprezentować swoje stanowisko oraz zaproponować działania mające na celu rozwiązanie problemu etycznego. Po upływie ustalonego czasu uczestnicy powinni omówić i zagłosować nad najefektywniejszym sposobem rozwiązania dylematu.

Rezultat: Ta aktywność fabularna umożliwi młodzieży rozważenie perspektyw licznych interesariuszy oraz krytyczne zastanowienie się nad odpowiedzialnością w systemach sztucznej inteligencji. Zachęci ich do refleksji nad złożonością odpowiedzialności w erze cyfrowej oraz nad tym, w jaki sposób różne strony przyczyniają się do podejmowania decyzji etycznych.

PRZECZYTAJ: Symulacje ról
| Centrum Innowacyjnego
Nauczania i Uczenia się



Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

4. Zaprojektuj swoją etyczną inteligencję sztuczną.

To kreatywne ćwiczenie umożliwia młodym ludziom refleksję nad procesem projektowania systemu AI od podstaw, uwzględniając jego etyczne implikacje. Angażując się w to praktyczne zadanie, uczestnicy zyskują lepsze zrozumienie, w jaki sposób zasady etyczne mogą być integrowane w procesie rozwoju.

Cel: Zainspirować młodzież do refleksji nad tym, jak można projektować systemy sztucznej inteligencji w sposób zgodny z zasadami etyki.

Instrukcje:

- **Formowanie zespołu:** Podziel uczestników na niewielkie grupy i przydziel każdemu z nich zadanie zaprojektowania systemu AI. Mogą oni wybrać konkretne zastosowanie, takie jak AI w rekrutacji, diagnostyce medycznej lub zarządzaniu inteligentnym miastem.
- **Etyczne projektowanie:** Każdy zespół powinien uwzględnić zasady etyczne w swoim projekcie. Należy omówić i rozwiązać kwestie takie jak uczciwość, przejrzystość, prywatność danych oraz inkluzyjność.
- **Prezentacja:** Po zaprojektowaniu systemu AI, każdy zespół przedstawi swój system oraz wyjaśni, w jaki sposób uwzględnił zasady etyczne w projekcie. Powinni również omówić potencjalne wyzwania etyczne, z jakimi może się zmierzyć ich system AI, oraz metody ich rozwiązania.

Rezultat: Ta aktywność rozwija kreatywność oraz myślenie etyczne, umożliwiając młodym ludziom aktywny udział w projektowaniu systemów AI, które stawiają na pierwszym miejscu kwestie etyczne. Pomaga im zrozumieć, że etyczne projektowanie AI nie jest sprawą drugorzędną, lecz kluczowym elementem procesu rozwoju.



CZYTAJ: Sztuczna inteligencja - przykłady dylematów etycznych.

Prowadzenie młodzieży poprzez dyskusje na temat etycznej sztucznej inteligencji.

5. Quiz etyki AI

Zabawnym i interaktywnym sposobem na sprawdzenie oraz utrwalenie wiedzy młodych ludzi na temat etyki sztucznej inteligencji jest quiz. Może być on przeprowadzany indywidualnie lub zespołowo, co zachęca młodzież do refleksji nad kluczowymi koncepcjami w lekki, lecz edukacyjny sposób.

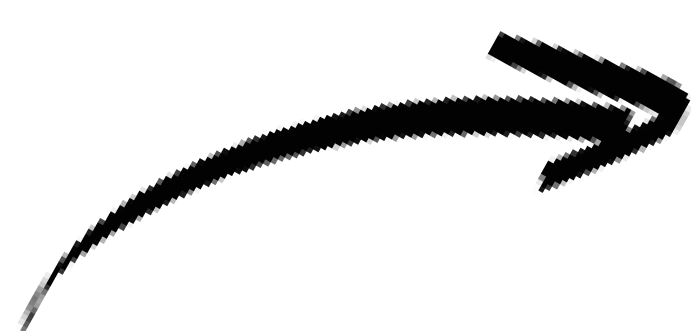
Cel: Ocena stopnia zrozumienia etyki sztucznej inteligencji przez młodzież w formie zabawy i aktywnego uczestnictwa.

Instrukcje:

- Przygotuj pytania: Opracuj zestaw pytań wielokrotnego wyboru lub pytań otwartych, opartych na treściach omawianych w trakcie dyskusji i ćwiczeń. Pytania mogą dotyczyć takich zagadnień, jak:
 - Czym jest stronniczość algorytmiczna?
 - Dlaczego przejrzystość jest istotna w systemach AI?
 - Jakie potencjalne zagrożenia wiążą się z zastosowaniem sztucznej inteligencji w procesie podejmowania decyzji?
 - Kto jest odpowiedzialny za szkody spowodowane przez systemy sztucznej inteligencji?
- Przeprowadź quiz: Zadaj pytania na głos lub wyświetl je na ekranie. Uczestnicy mogą odpowiadać indywidualnie lub w grupach, a poprawne odpowiedzi powinny być omówione po każdym pytaniu.

Rezultat: Ten quiz wspiera proces uczenia się i umożliwia młodzieży ocenę swojej wiedzy. To doskonały sposób na pobudzenie refleksji oraz utrwalenie zrozumienia kluczowych koncepcji etycznych związanych ze sztuczną inteligencją.

PRZECZYTAJ: Interaktywny quiz etyczny na temat sztucznej inteligencji



Jak prowadzić zajęcia dotyczące sztucznej inteligencji i etyki

Krótką, praktyczną listą kontrolną dla nauczycieli i pracowników młodzieżowych.

1 Przygotuj teren

- Zapoznaj się z fundamentalnymi zasadami etycznymi, takimi jak przejrzystość, uczciwość, odpowiedzialność oraz ochrona danych.
- Przejrzyj odpowiednie ramy (ustawa UE dotycząca sztucznej inteligencji, RODO, etyka sztucznej inteligencji UNESCO).
- Oceń umiejętności cyfrowe swojej grupy oraz jej wcześniejszą wiedzę na temat sztucznej inteligencji.
- Wybieraj studia przypadków odpowiednie do wieku oraz kontekstu kulturowego.

2 Zaprojektuj sesję, która będzie inkluzywna i angażująca.

- Używaj przykładów z życia (np. stroniczne narzędzia rekrutacyjne, rozpoznawanie twarzy, halucynacje w chatbotach).
- Zachowaj równowagę pomiędzy kwestiami technicznymi a dyskusjami na temat wpływu społecznego.
- Zachęcaj do krytycznego myślenia, stawiając sobie pytania takie jak:
„Czy sztuczna inteligencja powinna podejmować decyzje dotyczące przyznawania pracy?”
„Czy wykorzystywanie sztucznej inteligencji w edukacji lub policji jest etyczne?”
- Uwzględniaj różnorodne perspektywy i wspieraj otwarty dialog.

3 Korzystaj z interaktywnych metod dydaktycznych.

- Włącz multimedia (filmy, infografiki, symulacje).
- Korzystaj z interaktywnych narzędzi (np. demonstracji AI Fairness 360, narzędzia What-If od Google).
- Rozwiązuj dylematy etyczne lub uczestnicz w debatach („Jesteś programistą... czym się zajmujesz?”).
- Zastosuj krótki quiz lub ankietę, aby zainicjować refleksję i zwiększyć zaangażowanie.

4 Skupienie na wartościach i odpowiedzialności

- Połącz lekcje dotyczące obywatelstwa cyfrowego, równości oraz zrównoważonego rozwoju.
- Omów wybory ludzkie związane ze sztuczną inteligencją — selekcja danych, decyzje projektowe, przewidywane rezultaty.
- Podkreśl, że „to, że możemy, nie oznacza, że powinniśmy”.
- Podkreśl znaczenie zróżnicowanych i inkluzywnych zespołów technologicznych w celu redukcji uprzedzeń.

5 Połącz AI z codziennym życiem oraz przyszłą karierą.

- Pokaż, w jaki sposób sztuczna inteligencja wpływa na narzędzia codziennego użytku, takie jak algorytmy TikToka, asystenci głosowi oraz automatyczne tłumaczenie.
- Poznaj etyczne role w rozwoju sztucznej inteligencji, takie jak doradca ds. polityki AI czy etyczny projektant UX.
- Zachęć uczniów do rozpoznawania zagrożeń etycznych w aplikacjach i narzędziach sztucznej inteligencji, z których korzystają.
- Pobudzaj ciekawość: „Jakie wyzwania rozwiązałbyś za pomocą etycznej sztucznej inteligencji?”

6 Oceń oraz przemyśl

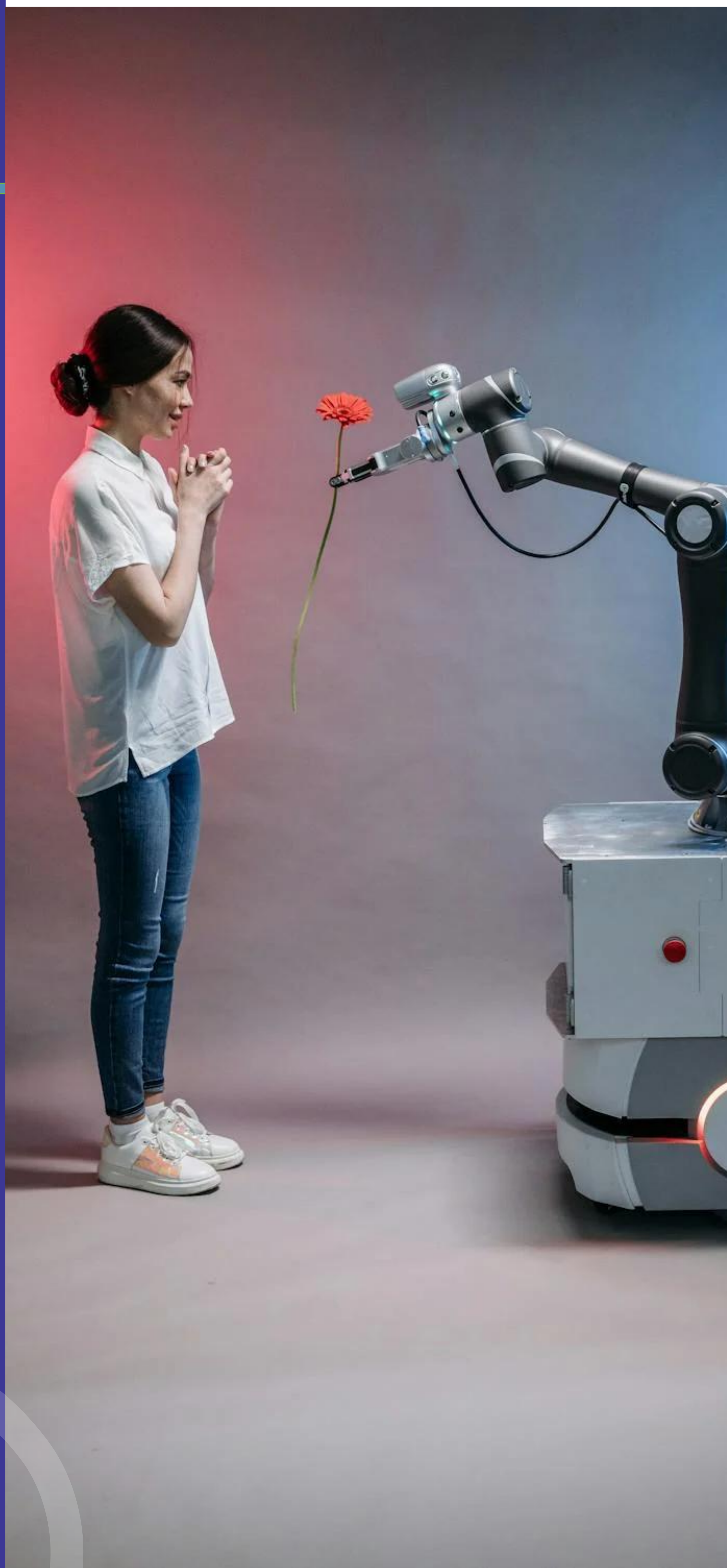
- Stosuj związane formularze z opiniami lub ćwiczenia refleksyjne.
- Zapytaj: „Czego dowiedziałeś się najbardziej?”
- Dostosuj przyszłe sesje w oparciu o zaangażowanie, zrozumienie oraz opinie.

Wnioski: Kierowanie młodzieżą w dyskusji na temat etycznej sztucznej inteligencji

W tym rozdziale zgłębiliśmy istotę prowadzenia młodych ludzi przez złożone zagadnienia etyczne związane ze sztuczną inteligencją. Ułatwiając dyskusje, zachęcając do krytycznego myślenia oraz wprowadzając interaktywne działania, młodzież może zrozumieć nie tylko mechanizmy działania sztucznej inteligencji, ale także jej wpływ na społeczeństwo oraz implikacje etyczne. Narzędzia te wspierają rozwój poczucia odpowiedzialności, umożliwiając młodym ludziom świadome zaangażowanie się w technologie sztucznej inteligencji i ich rolę w kształtowaniu przyszłości.

Działania przedstawione w tym rozdziale – od debat i odgrywania ról, przez studia przypadków, po ćwiczenia z kreatywnego projektowania – oferują edukatorom oraz pracownikom młodzieżowym praktyczne metody inspirowania świadomego, otwartego dialogu na temat etyki sztucznej inteligencji. Stanowią one zarówno edukacyjne, jak i transformacyjne doświadczenia, które zachęcają młodych ludzi do krytycznej analizy technologii, z których korzystają oraz które napotykają w codziennym życiu.

Pomagając młodzieży w nawigacji przez te dylematy etyczne, wyposażamy ją w umiejętności podejmowania odpowiedzialnych decyzji oraz angażujemy się w dyskusje na temat roli, jaką sztuczna inteligencja odegra w kształtowaniu ich życia, kariery i całego społeczeństwa. Niniejszy rozdział stanowi wezwanie do działania dla edukatorów i pracowników młodzieżowych, aby stali się moderatorami tych istotnych rozmów, prowadząc młodzież do stania się świadomymi etycznie i aktywnymi uczestnikami w świecie jutra, napędzanym przez sztuczną inteligencję.



„Kluczem do zrozumienia sztucznej inteligencji jest nie tylko wiedza o tym, jak ona działa, ale także zrozumienie, co oznacza ona dla ludzkości.”

Timnit Gebru,
Badacz sztucznej inteligencji i orędownik etycznego rozwoju sztucznej inteligencji.



05

ZAŁĄCZNIKI



Załączniki

Bibliografia

1. **Binns, R. (2018).** *"On the Logic of Trustworthy AI."* Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.
Retrieved from: <https://dl.acm.org/doi/10.1145/3173574.3173793>
2. **Bryson, J. (2018).** *"The Ethics of Artificial Intelligence: An Overview."* Oxford Handbook of Ethics of AI. Oxford University Press.
3. **Dastin, J. (2018).** *"Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women."* Reuters.
Retrieved from: <https://www.reuters.com/article/us-amazon-com-jobs-insight-idUSKCN1MK08G>
4. **European Commission (2020).** *"White Paper on Artificial Intelligence – A European Approach to Excellence and Trust."* European Commission. Retrieved from: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf
5. **Floridi, L. (2018).** *"The Ethics of Artificial Intelligence."* The Cambridge Handbook of Information and Computer Ethics. Cambridge University Press.
6. **Gebru, T. (2020).** *"AI and the Future of Humanity."* The New York Times.
Retrieved from: <https://www.nytimes.com/2020/09/03/technology/ai-debate-ethical.html>
7. **Gillespie, T. (2018).** *"The Relevance of Algorithms in Social Media."* Media, Culture, & Society.
Retrieved from: <https://journals.sagepub.com/doi/10.1177/0163443718768324>
8. **Goodall, J. (2019).** *"Ethics and Accountability in AI Systems."* Stanford Encyclopedia of Philosophy.
Retrieved from: <https://plato.stanford.edu/entries/ethics-ai/>
9. **O'Neil, C. (2016).** *"Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy."* Crown Publishing Group.
10. **Raji, I. D., & Buolamwini, J. (2019).** *"Actionable Auditing: Investigating the Impact of Publicly Available Face Dataset on Racialized Demographic Groups."* Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.
Retrieved from: <https://dl.acm.org/doi/10.1145/3293663.3293674>
11. **Russell, S., & Norvig, P. (2016).** *"Artificial Intelligence: A Modern Approach."* 3rd Edition. Pearson.
12. **Simonite, T. (2018).** *"How Deep Learning Algorithms Are Tackling AI's Bias Problem."* MIT Technology Review.
Retrieved from: <https://www.technologyreview.com/2018/02/26/65985/how-deep-learning-algorithms-are-tackling-ai-bias-problem/>
13. **Vincent, J. (2021).** *"What Are Deep Fakes and How Do They Work?"* The Verge.
Retrieved from: <https://www.theverge.com/2021/1/20/22242362/deepfake-ai-fake-video-deep-learning-explained>
14. **Zeng, J. (2020).** *"Transparency in AI and Ethics: The Importance of Open AI Systems."* Journal of Ethics and Information Technology.
Retrieved from: <https://link.springer.com/article/10.1007/s10676-020-09528-3>

Załączniki

Słowniczek

Sztuczna inteligencja (SI)

Dział informatyki koncentrujący się na tworzeniu maszyn lub oprogramowania zdolnych do realizacji zadań, które zazwyczaj wymagają ludzkiej inteligencji, takich jak percepcja wzrokowa, rozpoznawanie mowy, podejmowanie decyzji oraz tłumaczenie języków.

Algorytm

Zestaw zasad lub instrukcji, którymi posługuje się komputer w celu rozwiązania problemu lub wykonania zadania. W sztucznej inteligencji algorytmy wspierają maszyny w przetwarzaniu danych oraz podejmowaniu decyzji.

Błąd w sztucznej inteligencji.

Obecność systematycznej i niesprawiedliwej dyskryminacji w systemach sztucznej inteligencji, wynikająca z nie zrównoważonych danych lub wadliwego projektu algorytmicznego, prowadzi do niesprawiedliwych decyzji wobec niektórych grup.

Prywatność informacji

Praktyka ochrony danych osobowych polega na zapewnieniu, że są one wykorzystywane wyłącznie w sposób, na który dana osoba wyraziła zgodę. W kontekście sztucznej inteligencji, prywatność danych odnosi się do metod gromadzenia, przechowywania i wykorzystywania danych osobowych przez systemy sztucznej inteligencji.

Głębokie uczenie maszynowe

Podzbiór uczenia maszynowego, który obejmuje wielowarstwowe sieci neuronowe. Modele głębokiego uczenia wykazują szczególną skuteczność w zadaniach takich jak rozpoznawanie obrazów i mowy.

Etyczna sztuczna inteligencja

Systemy sztucznej inteligencji zaprojektowane, opracowane i wdrożone zgodnie z zasadami etycznymi, takimi jak uczciwość, odpowiedzialność, przejrzystość oraz poszanowanie prywatności.

Uczenie maszynowe (ML)

Rodzaj sztucznej inteligencji (AI), w którym algorytmy uczą się na podstawie danych i doskonalą swoją wydajność w czasie bez wyraźnego programowania. Jest on stosowany do zadań takich jak analiza predykcyjna, rozpoznawanie wzorców oraz wykrywanie anomalii.

Przetwarzanie języka naturalnego (NLP)

Dziedzina sztucznej inteligencji, której celem jest umożliwienie maszynom rozumienia, interpretowania oraz reagowania na język ludzki, stanowi fundament technologii takich jak rozpoznawanie mowy i chatboty.

Przejrzystość w dziedzinie sztucznej inteligencji

Zasada, że systemy sztucznej inteligencji powinny być przejrzyste i zrozumiałe, z wyraźnymi wyjaśnieniami dotyczącymi podejmowania decyzji oraz możliwością śledzenia i audytowania procesów, które leżą u podstaw wyników AI.

Odpowiedzialność w zakresie sztucznej inteligencji

Koncepcja, że twórcy oraz użytkownicy systemów AI powinni odpowiadać za wyniki działania tych systemów. W przypadku, gdy AI popełni błąd lub wyrządzi szkodę, istotne jest, aby ustalić, kto ponosi za to odpowiedzialność.

Zaangażowanie społeczne

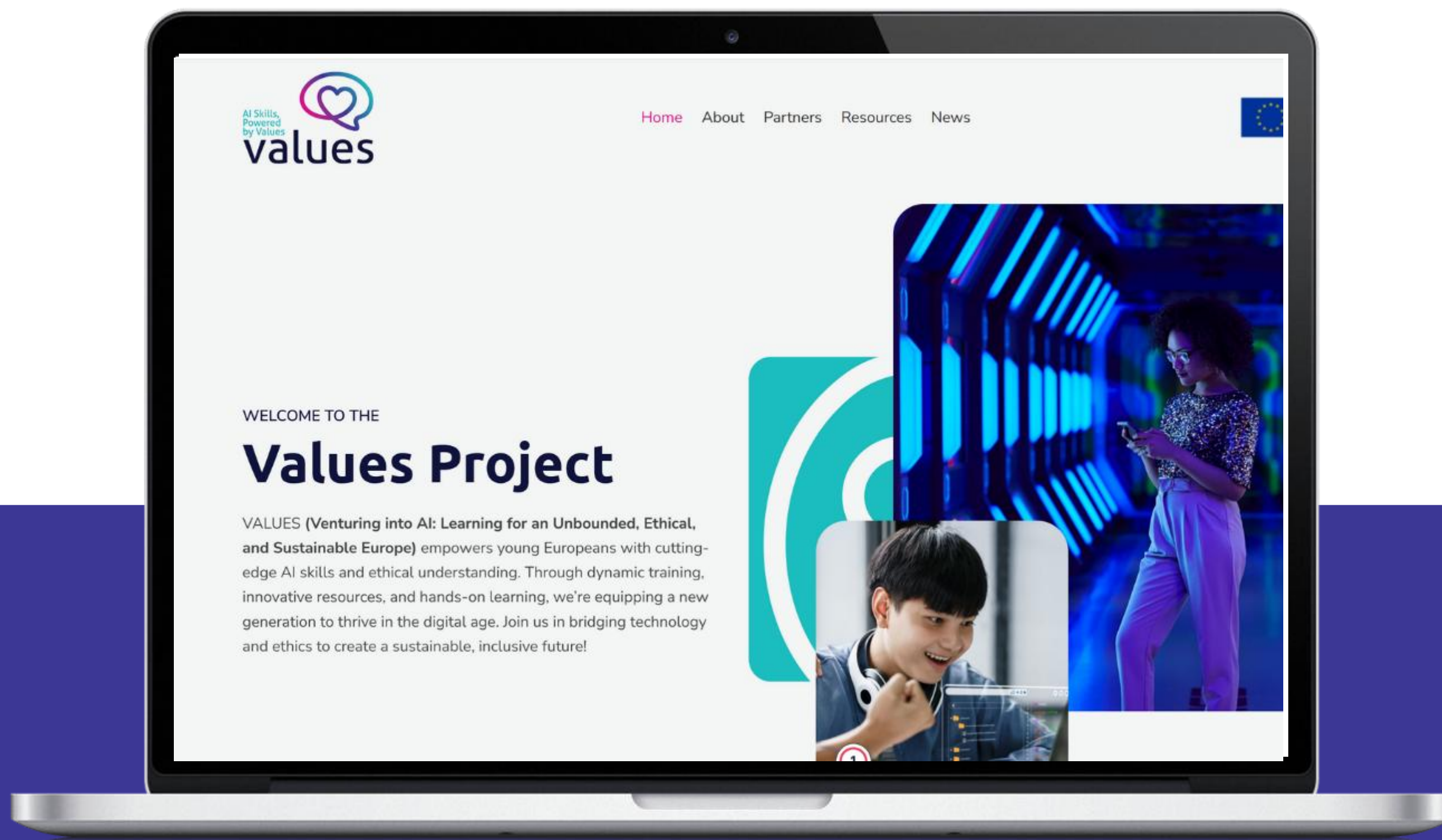
Zaangażowanie jednostek w społeczność, w tym uczestnictwo w procesach decyzyjnych oraz wykorzystanie technologii (takich jak sztuczna inteligencja) w celu odpowiedzialnego wspierania rozwoju społeczeństwa.

Sprawiedliwość w dziedzinie sztucznej inteligencji

Zapewnienie, że systemy sztucznej inteligencji traktują wszystkie osoby i grupy równo, bez faworyzowania ani dyskryminacji. Obejmuje to gwarancję, że błędy w danych lub modelu nie prowadzą do niesprawiedliwych wyników.

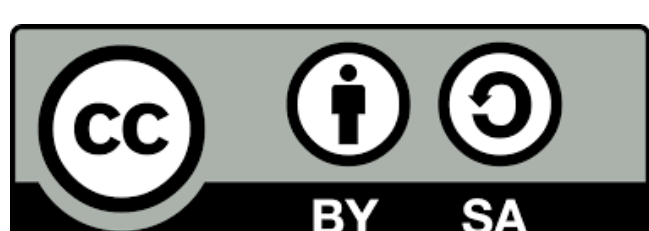
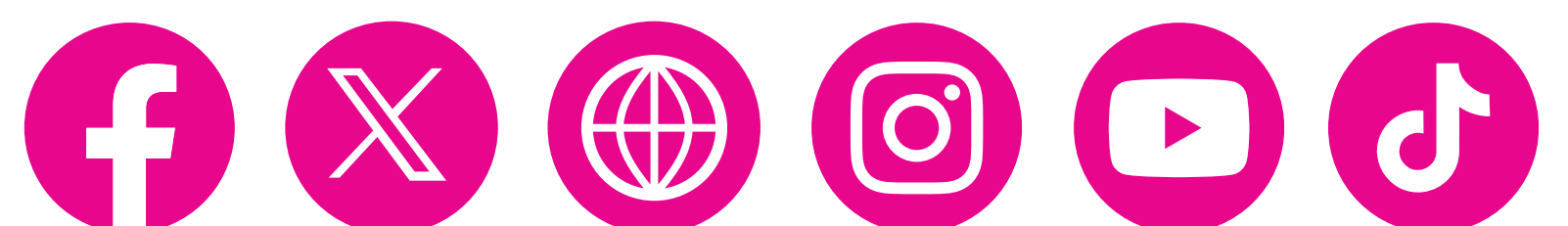
Odpowiedzialna inteligencja sztuczna

Rozwój i implementacja systemów AI, które gwarantują bezpieczeństwo, uczciwość oraz inkluzyjność. Obejmuje to zapewnienie, że aplikacje AI nie wyrządzają szkody jednostkom ani społeczeństwu oraz że spełniają standardy etyczne.



www.valuesai.eu

pozostań z nami



Co-funded by
the European Union